RESEARCH   Open Access

Group topic-author model for efficient discovery of latent social astroturfing groups in tourism domain

Noora Alallaq[1]* , Muhmmad Al-khiza'ay[1] and Xin Han[2]

Abstract

Astroturfing is a phenomenon in which sponsors of fake messages or reviews are masked because their intentions are not genuine. Astroturfing reviews are intentionally made to influence people to take decisions in favour of or against a target service or product or organization. The tourism sector being one of the sectors that is flourishing and witnessing unprecedented growth is affected by the activities of astroturfers. Astroturfing reviews can cause many problems to tourists who make decisions based on available online reviews. However, authentic and genuine reviews help people make informed decisions. In this paper a Latent Dirichlet Allocation (LDA) based Group Topic-Author model is proposed for efficient discovery of social astroturfing groups within the tourism domain. An algorithm named Astroturfing Group Topic Detection (AGTD) is defined for the implementation of the proposed model. The experimental results of this study revealed the utility of the proposed system for the discovery of social astroturfing groups within the tourism domain.

Keywords: Astroturfing, Astroturfing groups, Group topic-author model, Tourism

Introduction

The tourism sector is one of the fast-growing sectors which exists in many countries across the globe. People across the world plan their tours with the help of suggestions from others with prior experience. The new trend in decision making of tourists is the use of online reviews. The emergence of Web 2.0 technology features like micro-blogging gave came into existence, and this has paved way for online reviews in the tourism domain. Web sites like YELP.COM and Four-square.com facilitate online reviews that are reflective of the experiences of people that have previously used certain tourism facilities. Therefore, their opinions become very valuable to tourists who want to plan a trip. In this context, online reviews play the vital role of helping people make well informed decisions. However, fake reviews are a threat. The phenomenon in which a group of people or authors intentionally provide misleading online reviews through an organized campaign with the aim of promoting or demoting tourism services or packages is known as astroturfing. Astroturfing is carried out by an astroturfing group which is a set of authors. This is a challenge that must be addressed urgently. In spite of the fact that there are some mechanisms used for filtering the reviews of astroturfing, these mechanisms seem to be inadequate in terms of effectively filtering astroturfing reviews.

The literature (Shojaee et al., 2015; Rungta, 2015; Banerjee & Chua, 2015b) reveals that there are different models of detection astroturfing reviews. However, studies that focused on targeting astroturfing groups considering authors, documents and topics are limited in the literature. In this paper, we proposed a model known as Group Topic-Author model that is meant for the detection of topic-based latent tourism social astroturfing groups that are responsible for astroturfing campaign. For this study, tourism related online review datasets were collected from YELP.COM and Foursquare.com. The model is a generative probabilistic model that is based on LDA. There are three important distributions such as authors, topics and

* Correspondence: nalallaq@deakin.edu.au
[1]School of Information Technology, Faculty of Science, Engineering and Built Environment, Deakin University, Geelong, Australia
Full list of author information is available at the end of the article

hidden astroturfing groups characterized in the given dataset that is text corpora. The motivation for this work lies in the fact that the Group Topic-Author model is a novel approach which does not exist in the literature. The major contributions of this paper are as follows.

- The Group Topic-Author model for latent tourism social astroturfing group detection from online review corpus is proposed. This model is based on a modified LDA model which uses three parameters including, author distributions (a), latent astroturfing group distributions (β) and topic distributions (γ) in corpus.
- An algorithm that is named Latent Tourism Astroturfer Group Detection (LTAGT), which employs the use of unsupervised learning approach for clustering target astroturfing groups is also proposed.
- We collected 30 datasets containing online reviews from YELP.COM and Foursquare.com. Each dataset contains tourism reviews related to a restaurant.
- A prototype application is proposed to demonstrate proof of the concept. The experimental results of the study revealed the utility of the proposed methodology for building Group Topic-Author model.

The rest of the paper is structured as follows. Section 2 presents the reviews of literature on various approaches for filtering or identifying fake reviews given by online users. Section 3 presents the proposed methodology with author model, group topic-author model, and details about detecting astroturfing group by modelling the behaviour of astroturfing groups. Section 4 describes the experimental design used to evaluate the proposed methodology. Section 5 presents the experimental results for the detection of astroturfing group using group topic-author model, time and space complexity. In Section 6 the results are evaluated using performance metrics such as precision and recall. Section 7 presents discussion on threats to the validity of the proposed methodology. Section 8 draws conclusions and suggests directions for future work.

Related work
This section reviews literature on misleading online reviews and related research in tourism domain. It is divided into the following sub sections.

Mining and fake review detection on tourism websites
Tilly (Tilly & Cologne, 2015) studied online reviews and proposed a method for understanding user preferences in tourism domain. Again, in a study carried

out by Palumbo et al. (Palumbo & Rizzo, 2017) the location-based social networks for understanding the next stop and points of interests in tourism domain was studied. Banerjee and Chua (Banerjee & Chua, 2015b) made a textual analysis of tourism reviews related to hotels with the aim of distinguishing fake reviews from genuine ones. In their analysis, they were able to highlight the textual characteristics, which helped them in distinguishing the fake reviews from the genuine ones. In their study, More and Tidke (More & Tidke, 2015) proposed a weighting scheme and a framework for summarization of online reviews of tourism domain. Kumar et al. (Cardie & Hancock, 2011) studied consumer reviews and mined them in order to obtain rating for different products. Consequently, they developed a predictive model based on the ratings. Rungta (Rungta, 2015) studied online reviews with the aim of detecting spam in opinions given by tourists while making reviews. The dataset used by this author for experiments was obtained from TripAdvisor. Fong et al. (Fong et al., 2016) proposed a method of finding asymmetric of hotel rating within the tourism domain based on data obtained from TripAdvisor.

The reliability of reviews on TripAdvisor website was studied by Chua and Benerjee (Chua & Banerjee, 2013). They explored abnormalities by investigating highly inter-licked hotel. In a study conducted by Luca and Zervas (Luca & Zervas, 2015) it was found that suspicious reviews were on the increase in the tourism-related reviews available on YELP.COM. In another study carried out by Proserpio and Zervas (Proserpio & Zervas, 2016), it was found that the rate of fake online reviews detection on YELP.COM and TripAdvisor is on the increase as well. In order to estimate the prevalence of fake reviews on tourism websites like TripAdvisor, Priceline and Hotels.com, Ott et al. (Ott et al., 2012) proposed a generative model. Due to the increasing rate of fake reviews, Mukherjee et al. (Mukherjee et al., 2013) saw the need to study a tourism online review web site (YELP.COM) and its mechanism for preventing fake online reviews. These group of researchers opined that many review sites do not filter fake reviews.

Consumers' trust for online reviews is decreasing because of the increased evidence of fake reviews. In (Travellers Trust, 2015), the researchers investigated the trust of people on TripAdvisor, and found that web site quality, customer satisfaction and source credibility are crucial factors that influence the trust customers have for a website. Ott et al. (Cardie & Hancock, 2011) studied deceptive opinion spam in TripAdvisor and Yelp.com web sites. Dohse (Dohse, 2013) who investigated fake reviews on tourism websites, found that the boundary line between fake reviews and brand management is unclear. In order

to enhance the identification of deceptive reviews on tourism websites, Li et al. (Li et al., 2013) proposed a set of general rules. They were able to propose the general rules using the linguistic differences between fake and genuine reviews. in terms of spreading fake reviews to boost business visibility, Fisman (Fisman, 2012) found evidence of hotel owners spreading fake reviews to improve their popularity.

Spam detection on social media

Aichner and Jacob (Aichner & Jacob, 2014) focused on corporate social media and measured the degree of its existence and usage for garnering business intelligence. Banerjee et al. (Banerjee & Chua, 2015a) explored supervised learning algorithms for understanding and classifying online reviews into fake and real. Labbe et al. (Labbé et al., 2015) conducted a study using computer generated literature with the aim of detecting fake articles. Almagrabi et al. (Almagrabi & Malibari, 2015) carried out a survey related to product reviews and qualitative prediction of them. In order to identify sentiments in social media activities, Kumar and Sharma (Kumar & Sharma, 2017) analysed social media data. Bagnera and Suzanne (Bagnera & Suzanne, 2017) studied online rating of hotels to reveal the performance indicators of such hotels.

Kim et al. (Kim et al., 2017) introduced a novel database known as Paraphrase Opinion Spam as a learning mechanism for the accurate detection of opinion frauds. Pieper (Pieper, 2016) used Amazon to detect review spam, while Mahmood (Mahmood, 2017) explored and correlated journal rankings and their actual truth table.

Detection of astroturfing on online social networks

In order to detect hot topics in micro blogs, Ma et al. (Mat et al., 2014) proposed a topic model based on term correlation matrix. In their study, the term-topic matrix was obtained using Symmetric Non-Negative Matrix Factorization (SNMF). Nie et al. (Nie et al., 2017) integrated two approaches which are; topic model and word embedding to enable the clustering of suggestions from search engines. Their work focused on sub-topic clustering as part of text clustering.

Isupova et al. (Isupova et al., 2017) studied approaches for automatic analysis of behaviour. They achieved this using a learning mechanism for a topic model which they proposed. Sobolevsky et al. (Sobolevsky et al., n.d.) used data related to tweets, geotagged photographs, and bank card transactions for leveraging social media for foreign visitors. Cheng et al. (Cheng et

al., 2016) mined risk patterns from a database related to healthcare domain. Consequently, they proposed a model for the discovery of risk patterns using textual data. Shojaee et al. (Shojaee et al., 2015) proposed a framework for annotating fake reviews to support research related to the detection of astroturfing.

LDA-based approaches

Hassan et al. (Hassan et al., 2011) exploited multi-modal features for event detection from multimedia content. This solution is based on LDA model for topic modelling. Temporal and spatial combination with Conditional Random Fields is used to detect sport events. Huang et al. (Huang et al., 2014) proposed a topic model based on text clustering for important topic detection from micro blogs. Their method was an extension of the conventional Latent Dirichlet allocation (LDA). More so, Xu and Fan (Xu & Fan, 2015) employed multi-modal topic modelling for social event detection through geo-annotations. This is based on topic model derived from LDA. Similarly, Chen et al. (Chen et al., 2015) proposed a modified LDA model for the detection of popular topics on micro-blogs. With this, the reduction in capability related dimensionality was achieved. Zou et al. (Zou et al., 2016) also used topic modelling to identify duplicates in software bug reports. These researchers were able to overcome problems such as noise, sparse data and high dimension. In addition, they were able to achieve better performance as compared to the traditional Support Vector Machine (SVM). Through the use of LDA based approach, they were able to effectively identify topics in bug reports. Sendhikumar et al. (Sendhilkumar et al., 2017) employed the use of the concept of word clouds in order to have a topic modelling from text corpus in addition to LDA model. The aim of this was to generate topics and associated probabilities.

Proposed methodology for group topic-author model

This section provides details of the proposed group topic-author model meant for the discovery of tourism social astroturfing groups that are involved in spreading astroturfing reviews in tourism domain.

Problem definition

Online reviews have become crucial for decision making in the twenty-first century, because they are able to influence the decisions of people across the globe. With regards to products and services, online reviews

serve as means of obtaining information about the previous experiences of people with given product or service. This shows how valuable online reviews are in providing information about the merits and de-merits of a product or service. Thus, it is important to have such reviews to guide people on decisions related to the purchase of goods and services, or any other decision. However, the problem of astroturfing is associated with this. As found in the literature, astroturfing has been in existence for a while, and through it, people give misleading reviews with the sole aim of promoting or demoting services or products. Such fake reviews have become an issue of great concern. Astroturfing campaigns are carried out by a set of astroturfers (groups). In the literature many models have been introduced based on LDA. How-ever, a Group Topic-Author model aimed at discover-ing latent social astroturfer groups in tourism domain is lacking. This is the major problem addressed in this study.

In this paper, the problem is addressed by propos-ing a Group Topic-Author model that is based on LDA model and considers probabilistic distribution of authors, topics and latest astroturfing groups. For easy modelling, corpus of documents obtained from tourism domain are used in this study. Each docu-ment is nothing but an online review. Modelling Group Topic-Author model is non-trivial, as the model needs to reflect astroturfing behaviour of authors. The model facilitates the discovery of con-cealed latent astroturfing groups. The model con-siders authors, topics and also astroturfing groups behind the reviews. Here, it is important to have a time window since the astroturfing campaign lasts for just a certain period. The proposed generative model considers time window, documents, topics and hidden astroturfing groups.

Author model

The author model is meant to model authors and documents or reviews. It is a generative model which represents a set of authors and set of reviews. The LDA is not directly used for the aforementioned reason, instead, the variant of LDA used by Rosen-Zvi (Rosen-Zvi et al., 2003) which is provided in this sub section is used. We further improve it to make it Group-Author model as discussed in the next sub section. The model is provided to solely enable the modelling of authors' interests. The graphical representation of Author Model is shown in Fig. 1 below.

The boxes in the diagram are known as plates, and as seen in the diagram, there are many plates in this
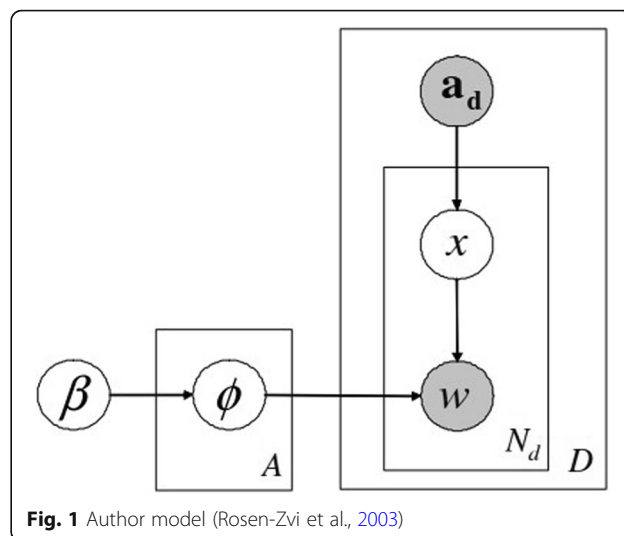

Fig. 1 Author model (Rosen-Zvi et al., 2003)

model. They represent replicates. The plate on the left hand side is a replicate of authors, the outer plate on the right side represents a set of documents, and the inner plate represents repeated choice of words within a document. Here, x refers to an author of a given word while ad indicates a set of users who pro-duced all the words. A probability distribution over words is denoted as θ, which is associated with each author. The probability distribution is generated from a symmetric Dirichlet prior denoted as β. Probability distribution has the potential of understanding author similarity. In spite of the benefits offered by this model, it is accompanied by some limitations. It is only able to provide author information as well as the words in their documents, and anything beyond that cannot be revealed the model. This potential limita-tion is overcome using the Group-Author Model proposed in this paper (Table 1).

Group topic-author model

This is also another variant of LDA which is genera-tive probabilistic model. In this model, documents are characterized based on the distribution. Topics that are distributed within the document are also consid-ered. Since it is a Group Topic-Author model, topics are also considered besides authors and documents. Latent astroturfers are also associated with documents that appear as random mixtures. Therefore, the Group-Topic-Author model is proposed as a novel approach to the discovery of latent tourism social astroturfers. The model needs LDA based approach that contains representation of documents in the

Table 1 Notations used in Group-Author model

Notation	Description
N_w	The number of words in a review.
N_p	The number of products in the dataset.
N_u	The number of reviewer in a single product.
G	The number of latent groups.
g	The group assignments for all reviewers.
θ	The G dimension group distribution.
α	The Dirichlet prior for group assignments distributions.
x	A reviewer of a product.
φ	Reviewer Distribution associated with a group.
β	The Dirichlet prior for word
l	The sentiment label of a review.
z	The sentiment distribution associated with a group.
π	The Dirichlet prior for sentiment distribution.
w	A word in a review.
γ	Words distribution associated with a sentiment label and a reviewer.
φ	The Dirichlet prior for words distribution.

Algorithm 1: Astroturfing Group Topic Detection (AGTD)

Input: Review Dataset D

Outputs: Initial Social Astroturfer Groups G

1: Initialize document corpus vector D
2: Initialize topic vector T
3: Initialize matrix of vectors V
4: Initialize Astroturfer group vector G
5: **for** each instance d in D **do**
6:     Extract review into a document d
7:     Add d to D
8: **end for**
9: Pre-processing
10: **for** each document d in D **do**
11:     Perform stop words on d
12:     Perform stemming on d using Porter Stemming algorithm
13: **end for**
14: Generating TF/IDF Matrix of Vectors
15: **for** each topic t in T **do**
16:     **for** each document d in D **do**
17:         **for** each word w in d **do**
18:             Generate TF/IDF matrix as vector v
19:         **end for**
20:     **end for**
21: Add v to V
22: **end for**
23: Finding Astroturfer Groups based on Topic
24: Apply Group Topic-Author model based on K-Means Clustering algorithm
25: Use cosine similarity to group vectors in V to form document clusters base on T
26: Group associated authors into social austrofuter groups G
27: Return G

outer plate and representation of words in the inner plate.

GTAM shown in Fig. 2 assumes the following generative process for each review R. The model first picks a group assignment g from a multinomial distribution θ. Then, according to the picked group g and a multinomial distribution ϕ, the reviewer x is generated. Meanwhile, for the picked group, a sentiment label l (positive or negative) for the target product is drawn from a multinomial distribution z. Afterwards, the generated reviewer x generates a review R according to the designed sentiment label l. As for the review R, each of its word w is independently drawn from a distribution γ defined by l and x. Finally, as a Bayesian generative model, we give each multinomial distribution θ, ϕ and γ, a prior distribution in the generative process. When we consider the complete model, the reviewer set is dynamically divided into G groups, each of which contains several reviewers. For each group, the genuine reviewer is allocated this group with a very lower probability. Then a predetermined threshold is used to filter the genuine reviewer. Therefore, the group only contains the reviewers which are most likely to belong to it after filtering. Based on the Group Topic-Author model, we defined an algorithm to detect astroturfing groups based on related topics. It takes tourism dataset as input and produces latent AGs which are associated with topics.

The LGTD algorithm initializes all the needed vectors (steps 1–4). Afterwards, all input reviews are extracted in the form of documents (step 5–8). In steps 9–12, pre-processing is carried out in terms of stop words and stemming. Steps 13–20, TF-IDF matrix is generated to represent words, topics and documents. In step 21, astroturfing groups are found based on topic by using the characteristics of astroturfing. K-Means algorithm alongside GTA model is used to generate document clusters associated temporal domain, and associated authors of each group are considered to be astroturfing groups (steps 21–24).

Algorithm 1 presents the process of discovering latent tourism social astroturfer groups using text clustering phenomenon. The algorithm takes 30 datasets containing tourism related reviews in Excel file format as input. The dataset contains attributes like author, review and date on which review is made by the author. The data presented in an Excel file format is subjected to text mining. For the sake of convenience, the dataset is converted into a document corpus denoted as D in the algorithm and denoted as α in the Group-Author model formally. Once the document corpus is ready, it is

Fig. 2 Proposed Group Topic-Author Model
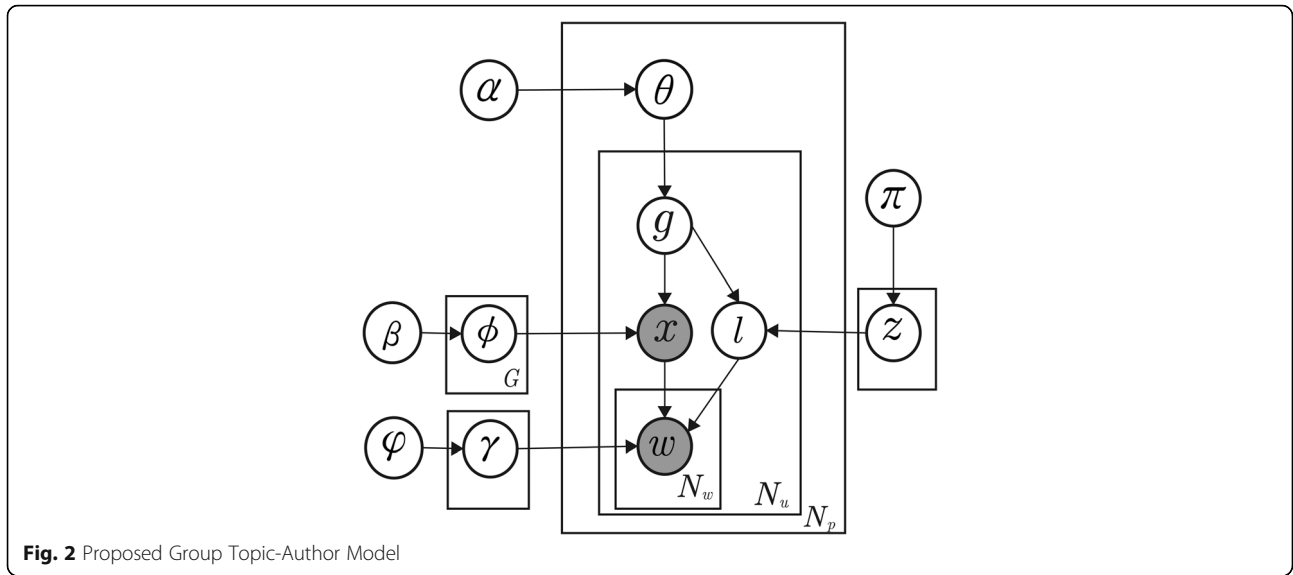
subjected to pre-processing, which occurs in two phases Pre-processing is made in two phases. In the first phase, the corpus is subjected to stop word removal. Stop words are the words in the set of documents (corpus) containing certain words that do not make any difference in the text clustering process.

After removing stop words, the corpus is ready for processing. However, before processing it is ideal for the corpus to go through stemming process which enables the identification of root words and removes all derived words. The well-known class PorterStemmer algorithm is reused here as a stemming mechanism. With stemming, the pre-processing ends. Now the documents in corpus are devoid of stop words and derived words. This implies that the corpus is ready for textual analysis. At this stage TF/IDF matrices are created one for each document base on given topics. TF/IDF stands for Term Frequency/Inverse Document Frequency. It is a standard measure used in reflecting the importance of a word to a document with respect to corpus. In fact, the vectors generated reflecting all documents contain information that can facilitate the process of clustering.

While performing clustering, we used Group Topic-Author model where three corpus level parameters are utilized. They are denoted as α, β and γ respectively. The former refers to author distribution in corpus while the second one denotes astroturfing group author distribution n corpus. The third parameter represents topic distributions. For the purpose of grouping, the Group Topic-Author model with K-Means algorithm is implemented. It generates clusters from collection of TF/IDF matrices that reflect tourism social astroturfing review clusters. The proposed Group Topic-Author model is able to group corresponding authors based on the associated clusters related to the given topics. The joint

probability distribution used in the proposed model is as follows.

$\theta|\alpha \sim Dirichlet(\alpha)$
$g|\theta \sim Multinomial(\theta)$
$x|g \sim Multinomial(\phi)$
$\phi|\beta : Dirichlet(\beta)$
$x|l : Multinomial(z)$
$z|\pi : Dirichlet(\pi)$
$w|x, l, \gamma : Multinomial(\gamma)$
$\gamma|\phi : Dirichlet(\phi)$

Solution Procedure

In the Group Detective Model, for a product, given the hyper parameters α, β, π and ϕ, and a set of N_u reviewers x_{ij}, the joint distribution of a reviewer mixture ϕ, a sentiment label l, a sentiment label mixture z, the R_{ij} presented by a set of N_w words w is given by:

$$p(R_{ij}, \theta, g, l, \phi, \gamma, z | \alpha, \beta, \phi, \phi, x_{ij}) = p(z|\pi)p(\phi|\beta)p(\gamma|\phi)$$
$$\times \prod_{j=1}^{N_u} \prod_{k=1}^{N_w} p(w_k|\gamma)p(x_{ij}|\phi)p(g|\theta)p(l|z)$$

By integrating over ϕ, γ, θ, z and summing over g and l, the marginal distribution of a review of a product is derived.

$$p(R_{ij}|\alpha, \beta, \phi, \phi, x_i) = \int \int \int \int p(z|\pi)p(\phi|\beta)p(\gamma|\phi)$$
$$\times \prod_{i=1}^{N_u} \prod_{k=1}^{N_w} \sum_g \sum_l p(w_k|\gamma)p(x_{ij}|\phi)p(g|\theta)p(l|z)\mathrm{d}\phi \mathrm{d}\gamma \mathrm{d}\theta \mathrm{d}z$$

Then we take the product of the marginal probabilities of single review, the probability of all the reviews of one product is:

$$p(R_i|\alpha, \beta, \phi, \phi, x_i) = \prod_{d=1}^{N_u} p(R_{ij}|\alpha, \beta, \phi, \phi, x_i)$$

Finally, we get the probability of all review of all product is:

$$p(R|\alpha, \beta, \phi, \phi, x) = \prod_{i=1}^{N_p} p(R_i)$$

At this point we use Gibbs sampling, a standard approximations method for graphical model (Sobolevsky et al., n.d.; Kumar & Sharma, 2017), to calculate.

$$p(g_t|x, R_i), t = 1, 2, 3, \cdots, G$$

This is the posterior distribution of the group assignments of a reviewer given the reviewer and their reviews. For each group g_t, t in 1, 2, 3, ..., G, the value of posterior distribution of g_t overall reviewers is calculated. Afterwards, a preselected threshold, eg. $k = 0.7$, facilitates the selection of the most likely reviewers who belong to the group g_t. As result, we get all members of each astroturfing group. Achieving this means that a solution to the problem earlier stated has been found.

Experimental setup

Experimental setup is required for the evaluation of the Group Author-Topic model that is proposed in this paper. For the experiments, 30 datasets related to tourism which were collected from YELP.COM (Yelp.com, 2017) were used. This datasets is available in Additional file 1. All the datasets are related to different restaurants associated with tourism. The restaurants include; are Aina, AracelyCafé, Barbacco, Beretta, Brendas French Soul Food, Burma Super Start, ChaChaCha, Coqueta, DermRestaurant, DumplingKitchen, DumplingTime, Fog Harbor Fish House, Frances, Francisca, Gary Danko, Hogwash, HoIsland Oyster Co, Hops And Hominy, HRD, Human, Izakaya Sozai, KElements BBQ, KuiShin Bo, LaFusion, LiholihoYachat Club, Little Skillet,

Lolo, MACD, Mano and Marlowe. JDK 1.8, Net Beans IDE are used as platform for implementing the proposed model and algorithm. A PC with 4 GB RAM and 1.70 GHz processing capability was used for the experiments. Performance metrics like execution time and memory consumption were used to determine the effectiveness of the proposed algorithm.

Experimental results

This section provides the results of the experimental results in terms of latent social astroturfing groups discovered from the datasets that were collected from tourism domain. As many as 30 datasets are used in experiments. The results are analysed and presented according to topics. The topics considered are food, price, location and service. Results also show performance metrics used such as, execution time and memory consumption for the proposed approach.

As presented in Fig. 3, it is evident that the topic "food" has more members in their groups. The dynamics of groups of all topics are visualized here. Five groups are associated with the four aforementioned four topics.

Listing 1: An excerpt from fake reviewers who are part of AGs
Topic: Food

Astroturfer Group 0: [Ziba Z, Pepatrip, priyanka141275, andreafi_166, Ben R., Mauri770304, Victoria A, Ali S, saleem m, Aitor K, Coneisha B, Abdullah A, Ashley Y.Los, Aishau89, San Francisco, mani k, Howie K.Albuquerque, Iris H.San Francisco, Xin W.San, Virginie D., Sakhar A, Victoria Y, Brian N.Oakland, Samiraahsharbi, Shahnawazthetraveler, Kara D.Twin, Giacomo G, Sandii M, Hossam G, RamblingGlutton, Mitzi G., Azza A, ColinsweeneY, Dianna H., Fairlady M, Julie L., sal l, Fathiya A, Julie C., JeffreyBlum, Lilly D., ConEfChriMa, westeam, kl61, marjane2011, silkwayhotel, rezashabanii, Meg W., Jae R.Sacramento, Bryan T., Dylan J, Amiromidi2017, Dean C., o Daryl-Blythe A.Fremont, Aathenaa, Joan M, Sam L., Vi T., E K, Venus L., Judy V., Tammy K., Rich F, Alisia B, Bob F., Loksanchari, Rayoody88, Rn T, ppiter, Sherry X.Berkeley, Mojtabashams, Richelle S, Marilyn T.Mountain, Winnie Y.Davis, Michalis C, Cew00, Sara A, kailuuu, MoeBhr, TPK751, Aishling H, Rob D, NicaNnewyork, Cotswolders, msn_8, adam456618, dorothy h, Dan B., cpw758, Michelle T.Orange, OrhanAlturk, Larissa C,
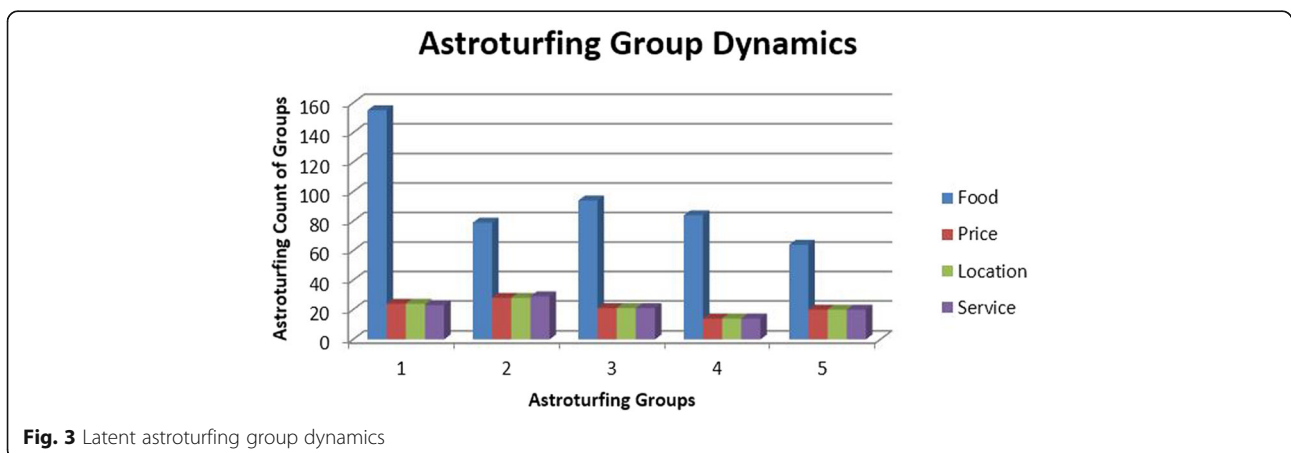

Fig. 3 Latent astroturfing group dynamics

Table 2 Summary of discovery astroturfing groups

Topic Name	Number Of Groups	Count of each Group				
		Group0	Group1	Group2	Group3	Group4
Food	5	155	79	94	84	64
Price	5	24	28	21	14	20
Location	5	24	28	21	14	20
Service	5	23	29	21	14	20

Experimental results *(Continued)*

NYandAussie, wxyz88, Mariani D., Ahmed S, Michalis, Kelly H, Sara-Anne M, Hamed K, Mahmoud I, Bryan S.Roseville, Alexander V, laatiif-faaa4566, Johnson N.San, Patrick C, Derby_Lad_200, fsolasf, Eng_haya, devilsdr, Sharon, Alireza B, Julienne, Raana B, Kamleish D, Dianna H.Daly, Inspector G, Dana B, Jane K, Haythamrs2, TheBob623, hindoya69, Suany W., Melanie N., Ken A, Paul W, sogoldavarzani06, RSG20140711, Jessica K, Abi A, Abby S.Denver, Sophie S., Richard G, Donald C., JKSang, TravellerG26, awolkiwis, JKBHerts, hhakim, yesim-tahmaz, Tracy D., qonieta, SSRH, Jenn R., ANDREAS V, Karen B, sben-harc, FAB1186338, Sheila S, Melissa K., K. P.Los Angeles, o Rhett B., Galliano16, Andrea K.San Francisco, Meliza M.Oakland, Candice D] Astroturfer Group 1: [Wendy L.Flushing, Michael C, Alladsprom, AaliyaeeeeD, Bianca P., Mei L., David G, Davi L, Mitzi G., Sarah M.San, eissahajji60, Mehak D.San Francisco, zen s, David K., Ikhlaq113, Kristen S., Patrik_Kerstin, E Z.San Marino, Gil S., Joanna D, Victoria, Ming Y.San Francisco, Val M., safy187, Jonathan N., Natascha E., Roddy S.Seattle, southernberry, Nimirta L, Jason W.Castro Valley, Daniel R., Mohannad_AlSharari, Shirley G., Sudipto G., Marie, Cecilia A.Newark, Chris Z.Berkeley, Joao S, Joseph L., A K., Daniel, Luci B.Queens, AASIF 133, Mehak D.San Francisco, Ichi Y., Jessie H.Manhattan, Jodie_L_Hart, Sheila H.Vienna, manaaarrrr65, Sikanderbakht, Eric Joseph D., Kim W., Tram N., Priscilla P.Morristown, Jason C.San, Hameed H, Andrea M., Sam C, Jean K., Iman-aoun, Don N., Lindsay, Andrey W, christinasmith2015, Ghazal S, Brigham, Abypune, Michael, Ian L.Los Angeles, Jeannie Z., Soundarya C., Chloe Anne, Justine J., DAEVA, Georges Albert, Geoff G.Santa, Meco P, Felipe L., DrPriyaS, JAZZ12, Franklin Z.Houston, arvindb2]

Here is an excerpt of fake reviewers that are part of AGs. It shows the astroturfing members belonging to a group. The summary of the groups is provided in the following sub section.

Summary of latest social astroturfing groups discovered in tourism domain

The summary of the astroturfing groups associated with different topics is provided in Table 2. The number of groups and count of each group are the two important items whose statistics is presented for each topic.

Five groups were considered for all the topics. More astroturfing group members were found in the case of groups discovered for the topic "food" when compared

Table 3 Performance of LTAGD algorithm in terms of execution time

Number Of Datasets	No. of Instances (all datasets)	Execution Time (Milliseconds)
30	6000	60204

Table 4 Performance of LTAGD in terms of memory consumption

Number Of Datasets	No. of Instances (in 30 datasets)	Memory Consumed (MB)
30	6000	85.43380737

with other topics such as price, location and service. The numbers of astroturfing group members in groups of the topic "food" are 155, 79, 94, 84 and 64 respectively. In case of the topic "service", the number of astroturfing members are 23, 29, 21, 14 and 20.

Execution time

The execution time of the proposed method is computed and presented in Table 3. The number of datasets, number of instances in all the datasets together and execution time in milliseconds are presented in the Table 3.

There are 6000 instances in the datasets. All instances are used in the experiments. The execution time of the algorithm LTAGD is recorded. It took 60,204 milliseconds to complete the execution of the proposed latent astroturfing group detection model.

Memory consumption

Memory is an important resource in computing machines. Memory consumption of the LTAGD algorithm is presented in Table 4. It shows the number of datasets involved in the experiments, the number of instances present in all the datasets, and memory consumed in megabytes. Memory consumed by the LTAGD algorithm is 85.43380737 MB. The consumed memory presented here is that which was consumed while processing the 6000 instances present in the30 datasets.

As presented in Table 4, it is evident that memory consumed by the LTAGD algorithm is 85.43380737 MB. The memory consumed is while processing the 6000 instances present in 30 datasets.

Evaluation of the proposed model

The Group Topic-Author model proposed in this paper was evaluated by comparing ground truth with experimental results. Industry experts were invited to evaluate the performance of the proposed model. The experts

Table 5 Confusion matrix used for evaluation

	Ground Truth (correct group decision for given topic)	Ground Truth (incorrect group decision for given topic)
Algorithm (correct group decision for given topic)	True Positive (TP)	False Positive (FP)
Algorithm (incorrect group decision for given topic)	False Negative (FN)	True Negative (TN)

Fig. 4 Precision-recall evaluation
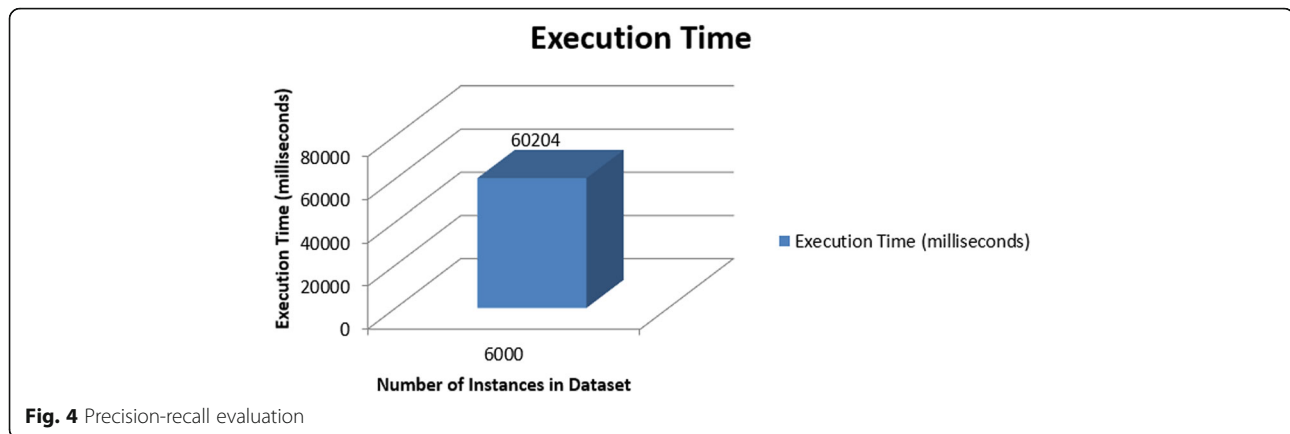
that evaluated the performance of the proposed model possess good domain knowledge and technical knowledge on astroturfing. They made ground truth with the help of the methodology provided with given datasets. Then they also used the proposed prototype application and datasets to compare the ground truth with experimental results. The ground truth values are compared with the result of the system for the purpose of evaluation. The results showed that the proposed system has its utility in detecting latent tourism social astroturfing groups. Evaluation is made based on the confusion matrix shown in Table 5.

We used two statistical measures for the evaluation, and they are known as precision and recall. The formal definition of those measure metrics shows as follow:

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

Precision and recall evaluation is visualized in Fig. 4. The evaluation results show the significance of the proposed Group Topic-Author model for the discovery of latent tourism astroturfing groups.

As presented in Fig. 4, it is evident that the recall and precision are presented in horizontal and vertical axes, respectively. According to the ground truth the proposed Group Topic-Author model with LTAGD algorithm showed high precision. It reveals the performance of the proposed model when it is evaluated according to the confusion matrix shown in Table 5. Based on the confusion matrix provided in Table 5, the precision and recall values are computed and presented in Fig. 4. As the precision increases, there is gradual decrease in the recall value. It is also evident in the graph that when recall increases, the precision is decreases.

Threats to validity

In this paper Group Topic-Author model is proposed for discovering latent tourism social astroturfing groups. The model is used with an algorithm named LRAGT) which takes tourism datasets as input an produces latent tourism social astroturfing groups. Datasets collected from YELP.COM are used with the prototype application to demonstrate proof of the concept. The results are evaluated by human experts with the help of ground truth. With respect to the evaluation results, there are threats to validity of the proposed methodology. The first threat to validity is that human experts considered for evaluation are very less in number, and that may not be sufficient to generalize findings. The ground truth provided by human experts might be biased. The third threat to validity is the usage of 30 datasets with 6000 combined instances. This dataset has limitations in terms of number of instances and coverage of tourism entities. Therefore, its correctness and generalisation may not be sufficient, thereby threatening the validity of the proposed methodology. The results could not be compared with other state-of-the-art model because no work on the detection of AGs was found.

Conclusions and future work

The tourism sector is a fast-growing sector in the world. The acquisition of information has been made easier with the emergence of micro-blogging, digitalization and usage of smart phones. People of all walks of life including tourists depend on online reviews while planning their trips, because online reviews can help them understand facts even before experiencing them as they were experienced by other people. Thus, online reviews are of great help to tourists and the tourism sector at large. However, astroturfing is one of the problems associated with online reviews. Astroturfing is a phenomenon in which misleading reviews are given by a group of individuals or astroturfers to influence the decisions of tourists. The review of literature revealed the fact that the

service providers of online review web sites do not have efficient mechanisms for filtering out astroturfing reviews. Therefore, in this paper we proposed an LDA based model known as Group Topic-Author model, and implemented an algorithm named Latent Tourism Astroturfing Group Detection (LTAGD) with unsupervised learning method to accurately identify astroturfing groups associated with astroturfing reviews. We also built a prototype application to show the efficiency of the proposed model. Tourism datasets from YELP.COM were collected as document corpus as input for the LTAGD algorithm. The experimental results revealed that the proposed model is useful in effectively identifying latent tourism social astroturfing groups. In the future, we intend to explore more on the proposed model with different domains and generalize its findings based on sentiment analysis.

Additional file

Additional file 1: Dada sets collected form Yelp.com. (RAR 10512 kb)

Availability of data and materials
Datasets used for experiments are available and they are in machine readable format. They are related to online reviews of tourism domain collected from YELP.COM. The datasets are provided in a separate compressed file.

Authors' contributions
NA has contributed in defining the proposed group topic-author model which is crucial for successful completion of the research. It includes the visual representation of generative probabilistic model that is used for implementation of prototype. MA-K on the other hand contributed in collecting datasets and moderate them to be machine readable and useful for experiments. XH contributed towards realization of the model with underlying mechanism which resulted in a prototype. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
[1]School of Information Technology, Faculty of Science, Engineering and Built Environment, Deakin University, Geelong, Australia. [2]College of Computer Science, Xi'an Shiyou University, Xi'an, China.

References
Aichner T, Jacob F (2014) Measuring the degree of corporate social media use. Int J Mark Res 57(2):1–19

Almagrabi H, Malibari A (2015) A survey of quality prediction of product reviews. Int J Adv Comput Sci Appl 6(11):1–10

Bagnera, Suzanne (2017) An examination of online ratings on hotel performance indicators: an analysis of the Boston hotel market. IEEE, pp 1–252

Banerjee S, Chua AYK (2015a) Using Supervised Learning to Classify Authentic and Fake Online Reviews. ACM, pp 1–8

Banerjee S, Chua AYK (2015b) Authentic versus Fictitious Online Reviews: A Textual Analysis across Luxury, Budget and Mid-Range Hotels. Information Science, pp 1–14

Chen Y, Li W, Guo W, Guo K (2015) Popular Topic Detection in Chinese Micro-Blog Based on the Modified LDA Model. IEEE, pp 1–6

Cheng Y-T, Lin Y-F, Chiang K-H, Tseng VS (2016) Mining disease sequential risk Patterns from Nationwide clinical databases for early assessment of chronic obstructive pulmonary disease. IEEE, pp 1–4

Chua AYK, Banerjee S (2013) Reliability of Reviews on the Internet: The Case of TripAdvisor, vol 1. WCECS, pp 1–5

Dohse KA (2013) Fabricating feedback: Blurring the line between brand management and bogus reviews. J Law Tech Policy, pp 363–393

Ray Fisman. (2012). Should You Trust Online Reviews, p1–4

Fong LHN, Lei SSI, Law R (2016) Asymmetry of hotel ratings on TripAdvisor: evidence from single- versus dual-valence reviews. IEEE, pp 1–38

Hassan E, Santanu C, Gopal M, Garg V (2011) A hybrid framework for event detection using multi-modal features. IEEE, pp 1–6

Huang S, Yang Y, Li H, Sun G (2014) Topic Detection from Microblog Based on Text Clustering and Topic Model Analysis. IEEE, pp 1–5

Isupova O, Kuzin D, Mihaylova L (2017) Learning Methods for Dynamic Topic Modeling in Automated Behavior Analysis. IEEE, pp 1–14

Kim S, Lee S, Park D, Kang J (2017) Constructing and Evaluating a Novel Crowdsourcing-based Paraphrased Opinion Spam Dataset. ACM, pp 1–10

Kumar N, Sharma A (2017) Sentimental analysis for political activities from social media data analytics. ICETETSM, pp 1–10

Labbé C, Labbé D, Portet F (2015) Detection of computer generated papers in scientific literature. ACM, pp 1–20

Li J, Ott M, Cardie C, Hovy E (2013) Towards a General Rule for Identifying Deceptive Opinion Spam, pp 1–11

Luca M, Zervas G (2015) Fake it till you make it: reputation, Competition, and Yelp Review Fraud, pp 1–35

Mahmood K (2017) Correlation Between Perception-Based Journal Rankings and the Journal Impact Factor (JIF): A Systematic Review and Meta-Analysis. Serials Review, pp 1–12

Mat H-F, Sunt Y-X, Jiat M-H-Z, Zhang Z-C (2014) Microblog hot topic detection based on topic model using term correlation matrix. IEEE, pp 1–5

More M, Tidke B (2015) A framework for summarization of online opinion using weighting scheme. Adv Comp Intel 2(3):1–9

Mukherjee A, Venkataraman V, Glance BLN (2013) What Yelp Fake Review Filter Might Be Doing, pp 1–10

Nie T, Ding Y, Zhao C, Lin Y, Utsuro T, Kawada Y (2017) Clustering search engine suggests by integrating a topic model and word Embeddings. IEEE, pp 1–6

Ott M, Cardie C, Hancock J (2012) Estimating the Prevalence of Deception in Online Review Communities, pp 1–10

Ott M, Choi Y, Cardie C, Hancock JT (2011) Finding deceptive opinion spam by any stretch of the imagination. Annual Meeting of the Association for Computational Linguistics, pp 309–319

Palumbo E, Rizzo G (2017) Predicting Your Next Stop-over from Location-based Social Network Data with Recurrent Neural Networks. ACM, pp 1–48

Pieper A-T (2016) Detecting review spam on Amazon with ReviewAlarm. ACM, pp 1–16

Proserpio D, Zervas G (2016) Online reputation management: Estimating the impact of management responses on consumer reviews, pp 1–43

Rosen-Zvi M, Griffiths T, Steyvers M, Smyth P (2003) The author-topic model for authors and documents. IEEE, pp 1–8

Rungta S (2015) Detection of opinion spam in online reviews. IEEE, pp 1–35

Sendhilkumar S, Srivani M, Mahalakshmi GS (2017) Generation of word clouds using document topic models. IEEE, pp 1–3

Somayeh Shojaee, Azreen Azman, Masrah Murad, Nurfadhlina Sharef and Nasir Sulaiman. (2015). A framework for fake review Annotation. IEEE, 1–6

Sobolevsky S, Bojic I, Belyi A, Sitko I, Hawelka B Scaling of city attractiveness for foreign visitors through big data of human economical and social media activity. IEEE, pp 1–10

Tilly R, Cologne (2015) An approach to derive user preferences from multiple-choice questions in online reviews. Twenty-Third European Conference on Information Systems, pp 1–11

Travellers Trust (2015) Why do travellers trust TripAdvisor? Antecedents of trust towards consumer-generated media and its influence on recommendation adoption and word of mouth, pp 1–43

Xu B, Fan G (2015) Multimodal topic modeling based Geo-annotation for social event detection in large photo collections. IEEE, pp 1–5

Yelp.com (2017). Retrieved from https://www.yelp.com/sf

Zou J, Xua L, Yanga M, Yan M, Yang D, Zhang X (2016) Duplication Detection for Software Bug Reports based on Topic Model. IEEE, pp 1–6