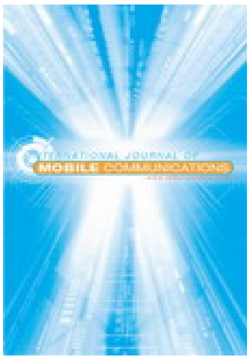

International Journal of Mobile Communications

ISSN online: 1741-5217 - ISSN print: 1470-949X
https://www.inderscience.com/ijmc

Predicting eBook purchases of heterogeneous social groups in a social network site using network metrics

Jongtae Yu, Dong-Yop Oh, Triss Ashton, Yang Wang

DOI: 10.1504/IJMC.2023.10042449

Article History:

Received:	02 February 2021
Accepted:	03 August 2021
Published online:	04 July 2023

Predicting eBook purchases of heterogeneous social groups in a social network site using network metrics

Jongtae Yu

King Fahd University of Petroleum and Minerals (KFUPM),
Academic Belt Road, Dhahran 31261,
Kingdom of Saudi Arabia
Email: jongtae.yu@kfupm.edu.sa

Dong-Yop Oh*

Department of Information Systems,
Auburn University at Montgomery,
7400 East Dr, Montgomery, AL 36117, USA
Email: doh1@aum.edu
*Corresponding author

Triss Ashton

Management Department,
Tarleton State University,
Box T-0330, Stephenville, TX 76402, USA
Email: ashton@tarleton.edu

Yang Wang

Information Technology Management,
University of Wisconsin-Milwaukee,
Milwaukee, WI 53211, USA
Email: wang400@uwm.edu

Abstract: This study examines users' social influence on e-book purchases within a social network drawing on the structural equivalence model. Structural equivalence holds that higher social influence levels exist among socially equivalent people (Burt, 1987). Using structural equivalence, network users were classified as either equivalent or inequivalent. Given that measurement data on social relationships among people within a network are often limited, to assign users to groups, link estimation utilised product choices to calculate network measures. With that framework, purchasing behaviours were predicted using various algorithms. Consistent with structural equivalence, the findings demonstrate that the average accuracy under the various algorithms is significantly higher in equivalent than inequivalent networks. Finally, comparing results with and without the network measurement variables suggests that failing to consider social equivalence may mislead prediction results by overestimating the social influence effect in low equivalent groups or underestimating the effect of high social equivalent groups.

Keywords: social influence; social network analytics; structural equivalence; classification; sales prediction; algorithm testing; biased predictions; network prediction performance.

Reference to this paper should be made as follows: Yu, J., Oh, D-Y., Ashton, T. and Wang, Y. (2023) 'Predicting eBook purchases of heterogeneous social groups in a social network site using network metrics', *Int. J. Mobile Communications*, Vol. 22, No. 1, pp.92–110.

Biographical notes: Jongtae Yu is an Assistant Professor at KFUPM Business School, King Fahd University of Petroleum and Minerals, Kingdom of Saudi Arabia. His primary research interests focus on online self-disclosure, privacy, and ethics. He has conducted studies to investigate online privacy concerns across different cultures and inconsistency between privacy concerns and disclosure behaviour. His work has been published in *Journal of Management Information Systems*, *International Journal of Healthcare Information Systems, and Informatics* and presented in several information systems conferences, including *INFORMS Conference on Information Systems and Technology*.

Dong-Yop Oh is an Associate Professor of Quantitative Methods at the College of Business, Auburn University at Montgomery. He received his master's and PhD in Applied Statistics from the University of Alabama. His research pertains to the areas of data mining, computational statistics, predictive models, and time series. He has opportunities to teach various courses, including business analytics, SAP PA, research methodology, and multivariate analysis.

Triss Ashton is an Associate Professor and Management Department Head at Tarleton State University. He holds a BS in Applied Physics from Angelo State University, an MBA in Decision Sciences, and a PhD in Management Science from the University of North Texas. He is a member of the US Society for Quality, where he is a Certified Manager of quality/organisational excellence. He has taught various courses, including business statistics, operations research, research methodology, and multivariate analysis. His research interests include text analytics, asymmetric distributions, quality control, and SERVQUAL.

Yang Wang is an Assistant Professor, Information Technology Management at University of Wisconsin-Milwaukee. His research focus is on the intersection of Marketing and Information Systems. One particularly interesting topic he is working on is that of product fit uncertainty in online markets. By collaborating with an online retailer, he aims to learn how different online artifacts help customers address the issue of product fit (mismatch). His work also addresses the issue of online network-related analysis, especially for the topics of purchase prediction using social network-based information. Yang received a bachelor's degree from Tongji University in 2010, and his master's degree in Marketing Research from Michigan State University and his PhD in Information Systems from University of Utah.

This paper is a revised and expanded version of a paper entitled 'Predicting eBook purchase in a social network site: structural equivalence model approach' presented at Decision Science Institute (DSI) in New Orleans, LA, USA, 23–25 November, 2019.

1 Introduction

The growing popularity of social media sites such as Facebook or Instagram offers a novel platform for e-commerce. Referred to as social commerce, people conduct commercial activities on these sites using social capital established in the social networking site (SNS) (Liang et al., 2011). Social media enables people to be active content creators, distribute information about products or services, and share their opinions, which significantly affect online purchases by others (Huang and Benyoucef, 2013).

Social networks as an e-commerce platform deserve greater attention. Globally, e-commerce was projected to exceed $26T in 2020, and with a 4% or higher annual growth rate (Lipsman, 2019) will exceed $30T by either 2024 or 2025. However, 2020 experienced a 24.1% increase in e-commerce sales to consumers, while total retail sales only increased 1% to $21.21T (Young, 2021). Discoveries in this unique marketplace can be applied to many other self-produced products or services. The market is fascinating as it provides new opportunities for small and medium-sized businesses, microentrepreneurial enterprises, and what Varian (2005) calls micro-multinationals. In the US alone, 99.9% of business are small, and they account for 47.5% of employment (US Small Business Administration, 2018); globally, micro, small, and medium-sized business make up 90% of the business and 50% of the employment (Worldbank, 2021). Many of those enterprises cannot afford to spend money on marketing or promotion. Beyond SNS e-commerce, the analysis procedures outlined here offer potential alternative forecasting strategies anytime collections of buyers or sellers purchase with varying velocity. Several global marketplaces meet that criteria.

Previous studies in social commerce empirically identified purchasing determinants on a social network site and their effects, including word-of-mouth (WOM) (Gunawan, and Huarng, 2015; See-To et al., 2014; Trusov et al., 2009; Wang et al., 2012; Wang and Yu, 2017); network externalities (Chiu et al., 2013; Katona et al., 2011; Xiao et al., 2018); or trust (Dabbous et al., 2020; Leeraphong and Mardjo, 2013; Lin et al., 2019b; Zhou, 2019). However, the increasing peer-to-peer transactions on social network sites require a greater social influence understanding for explaining a social entity's economic behaviour, such as product choice or adoption (Handarkho, 2020; Hu et al., 2019; Kietzmann et al., 2011). Since economic behaviours occur among the directly or indirectly connected users in a social network, a user can affect others' product purchases by sharing information about a product or by making recommendations through the established social relationship (Chu and Kim, 2011; Li and Lu, 2011; Li and Lai, 2014). This behaviour emphasises that social influence is exerted on economic behaviours in a social network at different levels (Aral and Walker, 2011; Liang et al., 2011; Lin and Lu, 2011; Ma et al., 2010; Rapp et al., 2013).

Social influence is also essential in predicting sales or transactions on social networks. Big data that capture transactions on social networks can enable predicting sales by analysing social influence, providing valuable insight to businesses (Chang and Li, 2019; Lee and Choeh, 2020). However, predicting sales in such environments is challenging. Predicting sales on social networks without considering the social users' heterogeneity is especially difficult. It will lead to highly biased solutions with select user groups overestimated while simultaneously underestimating other groups, as shown in this research.

The structural equivalence model (Burt, 1987) holds that structurally equivalent people occupy the same position in the social structure and interact with other individuals in similar ways. An observed structural equivalence occurrence is social contagion' manifestation, which leads to social homogeneity among persons (Burt, 1987; Friedkin, 1984). Structurally equivalent persons show more homogeneity in their attitude or behaviour than non-equivalent positions (Harkola and Greve, 1995). That is, social contagion is more prevalent among people in the same social position.

This research draws on the structural equivalence model and attempts to predict e-book purchases on the SNS using a network analysis method. Toward that end, various network measures, such as network centrality, are calculated and utilised to predict e-book purchases. The network measures usage is particularly appropriate for analysing and predicting purchases when the direct associations measurement or estimation amid the products is not feasible. Thus, algorithms such as collaborative filtering (CF) and content-based methods are ill-suited due to the lack of data for estimating associations. Using the network measures, we predict the sales on a social network, using well-known classification algorithms AdaBoost, GA-ensemble, neural networks (NN), decision tree, random forest, and support vector machine (SVM). In addition, we examine the social equivalency effect by comparing sales predictions between a socially equivalent and an inequivalent group.

This research utilises data scrapped from an SNS in China exclusively designed for e-book sales and is analogous in functionality to eBooks.com or Selz in the US. The e-book marketplace selection to illustrate this research is because of the growth and shifts in marketing within that market. The growth is fuelled in part by technological developments that spur independent (indie) or self-publishing. In 2015, 727,125 ISBNs were assigned to self-published titles representing 625,327 individual self-published (indie) books (Anderson, 2016; BookBaby, 2017). Global e-books sales in 2018 were $13 B and predicted to reach $15.2 B by 2023 (Statista, 2019).

According to WalktheChat, a social media marketing agency, the mobile reading market in China had grown to more than 300 million monthly active users in 2017. Moreover, compared to the traditional market (e.g., print books or other material goods), consumers' e-book purchase behaviour and information sharing about their reading preferences are more coherently connected with the online social network. As such, the Chinese SNS offered us an apt context to address the relevant research questions.

In the China SNS, when an author uploads an e-book, information about the book is diffused to other users by the system, particularly to users linked to the author. That diffusion mechanism or the marketing support offered by service providers such as Post Planner helps build engagement with characteristics that facilitate identifying distinct social positions on the SNS. The engagement efforts drive network formation making the e-book marketplace ideal for this research. Authors and readers are all SNS users. They commonly share their opinions, photos, or videos similar to other SNSs. However, the authors and readers have different social positions as sellers and buyers. This research suggests that authors are a socially equivalent group due to their shared interest in writing and their jobs as authors and e-books sellers; whereas, readers emerge from all walks of life and are socially inequivalent.

This research attempt to answer two fundamental questions:

1 Can we accurately predict e-book sales in the social network using network measures that capture social influence?

2 The prediction from authors (socially equivalent group) more accurate than that from readers (socially inequivalent group)?

This study contributes to the literature in several ways. First, it suggests an alternative way of using product purchases as nodes for structural measurement when accessing data on social relationships is limited. Second, it demonstrates that a person's product purchasing has a greater effect on others' purchases when their position in the social network is equivalent. Third, the findings suggest that the network approach is helpful in forecasting transactions in SNSs. Greater prediction accuracy occurs in the high social equivalent groups. As a further contribution, comparing results with and without structural equivalence suggests that failure to consider social equivalence may mislead predictions by overestimating low social equivalent groups or underestimating high social equivalent groups influence. Collectively, the results suggest that SNS-based predictions need to go beyond the traditional prediction variable set and include network measures.

2 Literature review of related works

Social network studies have taken many approaches. Hromic and Hayes (2019) address community detection by describing substructures in microblogging as either functional or structural communities. Functional communities form around common yet independent social interests or activities. Hromic and Hayes (2019) offer football team fans as a functional community example. In contrast, structural community membership depends on connectivity in the network that can be measured by the average node degree.

Beyond community detection, previous studies in social commerce empirically identified purchasing determinants on an SNS and their effects across the network by features such as WOM information dissemination (Gunawan and Huarng, 2015; Lu et al., 2016; See-To and Ho, 2014); network externalities (Chiu et al., 2013; Katona et al., 2011; Teh et al., 2015); or trust (Hajli et al., 2017; Leeraphong and Mardjo, 2013).

Lu et al. (2016) examine three social presence aspect effects (social presence in online environments, perception of others, and social interaction with sellers) on purchase intention via trust in sellers. They found a significant positive effect by the social presence factors on trust, leading to purchase behaviours. Hajli (2014) also shows that social commercial application encourages social interactions by increasing trust, which, in turn, fosters purchase intention. Chen and Shen (2015) demonstrate that emotional and informational social support is positively associated with trust and community commitment, affecting social shopping and social sharing intention.

Beyond trust, the increase in peer-to-peer transactions on SNSs requires a greater social influence understanding for explaining a social entity's economic behaviour, such as product choice or adoption (Kietzmann et al., 2011; Wang and Yu, 2017). Since economic behaviours occur among the directly or indirectly connected users in a social network, a user can affect others' product purchases by sharing information about a product or by making recommendations through the established social relationship (Li and Lai, 2014; Liu et al., 2016; Wang and Yu, 2017). Celebrity endorsements work in a similar manner (Zhu et al., 2020). This behaviour emphasises that social influence exerts pressure on economic behaviours in a social network at different levels (Aral and Walker, 2011; Liang et al., 2011; Lin and Lu, 2011; Ma et al., 2010).

Previous studies have predicted user behaviour by analysing social influence generated by frequent social interaction or possessing a similar position in a social network (Han and Park, 2020). For example, Bhatt et al. (2010) demonstrated that based on social pressure from friends within a social network, a user is more likely to purchase or adopt a product when his or her friends widely adopt the product. Guo et al. (2011) also suggest that people are more likely to purchase a product when their friends have already purchased from that vendor. In this circumstance, identifying a user who significantly influences other users in the social network is critical because influencers supply the standing necessary to convey a message to other users (Kiss and Bichler, 2008).

Wang and Yu (2017) examine the social influence on purchase behaviour by analysing WOM communications among users. They found that WOM communications content and valence affect purchase behaviour in an SNS. Similarly, electronic WOM (eWOM) systems studies in China positively affected behaviour intentions based principally on a reviewer's quality and argument quality (Goh et al., 2017). The reviewer's quality and argument quality are interpretable as the mechanism that transmits social influence.

Overall, our literature review suggests that many previous studies adopt an empirical approach for examining factors that affect purchase behaviour on social networks to explain the behaviour. However, surprisingly scant studies mine data and predict product or service sales in social networks, despite its usefulness in providing valuable insight to businesses. Moreover, although social influences' importance in explaining the behaviour is well acknowledged and examined, users' social equivalence on social networks is rarely considered. Ignoring social equivalence in analysing social influence for explaining or predicting a phenomenon may lead to biased results.

3 Hypotheses

A change in a person's cognition, attitude, or behaviour resulting from interaction with another individual or group defines social influence (Raven, 1964). Specifically, social contagion explains one entity's social influence that either an individual or group has on another. Social contagion refers to the social or psychological influence spread through direct or indirect interaction between individuals (Burt, 1987; Fenzl and Pelzmann, 2012). The social cohesion model suggests that social contagion occurs when parties have frequent and empathetic communication between them, which leads to engagement in similar behaviour (Burt, 1982). Social entities tend to exhibit similar behaviours or practices due to the evaluation and benefits of the shared cost associated with the activity or product under consideration. Thus, similar social entities demonstrate higher interaction rates than do dissimilar entities, which results in similar beliefs or values (Capozzi et al., 2016; Verbraken et al., 2014). In the SNS environment, an entity increases social influence as the entity has more connections with other entities or more frequent interactions with connections (Hudson et al., 2016; Molden and Dweck, 2006).

Prior studies often employ the social contagion model in explaining users' behaviour, including product purchasing in SNSs (Hinz et al., 2014; Hollenbaugh and Ferris, 2014; Yoo and Alavi, 2001). Fang et al. (2013) suggest that the probability of adopting a product or service is predictable by analysing social links based on social influence, structural equivalence, and entity similarity. However, while the empirical approach has been dominant in examining social contagion' effect, scant studies adopt a predictive analytic approach for forecasting users' economic behaviour on an SNS (Kim and Srivastava, 2007; Stephen and Toubia, 2010). Furthermore, analysing behaviours through the social contagion lens may introduce biases by ignoring the users' social heterogeneity. The social contagion model implicitly assumes no heterogeneity of social class or users' position in social interactions in a network.

A user's social position affects the social influence that occurs during interactions (Burt, 1987). The structural equivalence model holds that structurally equivalent people occupy the same position in the social structure and interact with other individuals in similar ways (Burt, 1987). An observed structural equivalence is social contagions' manifestation, leading to social homogeneity among persons (Burt, 1987; Friedkin, 1984). Structural equivalent persons show more homogeneity in their attitude or behaviour than non-equivalent (Harkola and Greve, 1995). In this light, social contagion is more prevalent among people in the same social position. Although social interactions are equivalent, social influence in the same social position would be higher than heterogeneous users. The structural equivalence model has been widely used to examine social influence in different settings (Burkhardt, 1994; Friedkin, 1984, 1993; Hinz et al., 2011; Pallotti and Lomi, 2011). Thus, the structural equivalence consideration would better predict users' economic behaviour in an SNS.

The social network approach has been applied for examining the effect of social influence and equivalence (Peng et al., 2016; Peng et al., 2017). According to social network theories, a network consists of nodes and edges representing individuals in a network and their relationships (Freeman, 1978). In this vein, the network approach emphasises the importance of relational data that exhibits how individuals are related to each other in a network (Krause et al., 2007). An individual's position is also an essential component that affects relationships in a network because individuals' interactions depend upon the social structure (Burt, 1987). Hence, network measures such as centrality or betweenness have often been used for analysing social influence in a network (Almgren and Lee, 2016; Chung et al., 2021). Therefore, the study hypothesises that:

Hypothesis 1 Using structural equivalence, two broad network classes (authors and readers), and their membership is discoverable in the SNS using network measures.

Hypothesis 2 Of the two network classes, authors will show more social interactions than readers.

Hypothesis 3 The author's network, a structural equivalent group, will demonstrate more substantial social influence on e-book purchasing in the SNS than the reader network, which is inequivalent.

4 Data

Data collection for this study utilised a web-crawler over an SNS in China. The SNS provides an online platform for e-book sellers, where each registered user can create e-books and sell them on the platform. Since it is also a social networking website, different users (including authors) connect, forming a unique network structure. Those e-books are sold exclusively on the social network platform; therefore, exogenous factors that could affect the purchase decision are very limited.

The data contains 530,470 e-book purchase transaction records for approximately 6,000 e-books sold on the SNS. By examining the offerings genre, the data was reduced and focused. Of the 6,000 e-books, only 2,021 titles include genre information in the transaction records. By selecting the novel (fiction) genre produced a dataset that accounted for approximately 61% of all e-books with genre information. After cleaning the data and removing duplicate transactions, the final dataset holds 78,896 sales transactions.

The data comprises three scraped data files. The transaction records include the users' identification code, transaction date, e-book identification code, recommender rating (1–5 if present, else 0), rating date, and transaction type (purchase, subscribe, or gift). A data sample is included in Table 1. The recommendation records identify books purchased and recommend additional titles based on prior purchase behaviour, customer ratings, and activity date. A data sample is available in Table 2.

Table 1 Sample of transaction records

User_ID	*Date*	*Book_ID*	*Rate*	*Rate_time*	*Behaviour*
1119306	6/17/2015	363081	5	9/24/2014	purchase
3430534	6/17/2015	11279632	4	9/14/2015	gift
1938286	6/17/2015	7533672	0	None	purchase
1994295	6/17/2015	1667174	0	None	subscribe

Table 2 Sample of recommender data

Purchased book_Id	*Recommended_book_Id*	*Date*
1811	340587	6/18/2015
1811	381612	6/18/2015
2432	440810	6/18/2015
2432	686991	6/18/2015

The final data file records the book information. That file consists of the e-book ID, author ID, book type, and word count (rounded to thousands) (Table 3).

Table 3 e-Book information

Book_Id	*Author_Id*	*Type*	*Word_count*
1420957	63689185	novel/novella	19,000
2610048	63693845	comics/novella	None
4683406	63695116	nonfiction/novella-collections	4,000
1782989	63688078	novel/novella-collections	25,000
2848915	63693719	photography/long-story	None

From the data, two networks (author and reader) based on users' social positions on the SNS are distinguishable and used to examine the social influence by structural equivalence. The author network comprises users who upload e-books to sell to other users. The authors constitute a unique social group on the SNS with similar motivations and interests. In contrast, the reader network users buy e-books. The users' social position in this group is inequivalent because they possess various social positions, interests, or values on the site. The data includes 23,447 transactions between authors and 55,449 transactions between authors and non-authors (i.e., readers). The scrapped data covered four years. For analysis, model training utilised the first two years of data, with the remaining two years of data reserved for testing. Table 4 offers the sample sizes used for training and testing purposes.

Table 4 Data sample sizes used for training and test (unit: transactions)

	Training	*Testing*	*Total*
Author network	10,137	13,310	23,447
Reader network	38,040	17,409	55,449

The data is limited in that information about social links or interactions among the entities in the SNS was missing, which is commonly the case (Cha et al., 2010; Kumar et al., 2010). Therefore, the social link estimation utilised a product network approach. The product network approach analyses the economic entities behaviour based on the assumption that entities establish social links when they have similar preferences toward a product or service (Dhar et al., 2014). A similar preference toward a product or service offers a foundation for establishing relationships among people. Social entities in a social network with similar preferences toward a product will find it easier to establish a relationship (Fang et al., 2013). In a product network, products are the network nodes instead of the social entities, and shared economic outcomes offer a foundation for establishing the links between the social entities (Dhar et al., 2014). For example, two simultaneous product purchases signal the existence of the social link. This approach facilitates an alternative means of establishing connections between social entities. Two entities are acknowledged as linked on a social network when they frequently buy the same product. Structural measures were calculated with social links using the Gephi™ software package. The specific network measures computed are defined in Table 5.

A network measures summary for the two networks is available in Table 6. The results demonstrate that consistent with hypothesis one, two distinct networks exist in the data, and consistent with hypothesis two, the authors possess more social links than readers. That suggests more social ties because they are structurally equivalent users.

Table 5 Computed network measures

Network measure	*Definition*
Degree	The total number of relations (edges) that a node has
In-degree	Incoming edges count
Out-degree	Outgoing edges count
Weighted degree	A weighted sum of the number of edges
Modularity-class	How well a network decomposes into modular communities/clusters
Eccentricity	The distance from a node to the node that is the farthest away
Closeness centrality	Average shortest path length between a node and other nodes
Betweenness centrality	The number of times a node acts as a bridge between other nodes
Clustering coefficient	The average clustering coefficient of all nodes in the network
Eigenvector centrality	An influence measure for the node on the network

Table 6 Calculated author and reader network measures

	Author network		*Reader network*	
	Training	*Testing*	*Training*	*Testing*
Average degree	36.596	25.132	11.584	13.283
Avg. weighted degree	18.298	12.566	5.819	6.677
Graph density	0.033	0.025	0.001	0.001
Modularity	0.837	0.718	0.943	0.906
Avg. clustering coefficient	0.484	0.472	0.396	0.375
Avg. path length	1.557	2.642	3.984	5.537

5 Data analysis and results

Purchase prediction by users began by distinguishing between the two network types (author and reader), modelling each network, then testing the resulting model to see which network has higher predictive accuracy. Model development used the well-known classification algorithms AdaBoost (Freund and Schapire, 1997), decision tree (Breiman et al., 1984), GA-ensemble (Oh and Gray, 2013), neural network (McCulloch and Walter, 1943), random forest (Breiman, 2001), and SVM (Cortes and Vapnik, 1995). Comparisons concentrate on accuracy and recall as performance measures to balance the evaluation of the analytical approach. Consistent with traditional algorithm comparison practices, reporting includes precision and F-measure statistics. Accuracy and recall are the central focus because, to the data provider, as well as the booksellers, it is more important to predict the buyers accurately [(true-positive + true-negative) / (true-positive + true-negative + false-positive + false-negative)], and recall [true-positive / (true-positive + false-negative)], the relevant instances predicted proportion. Wrong predictions are not a major concern because the low cost of targeting potential buyers is negligible. As long as the potential buyer is a follower in this SNS, sellers can quickly send them promotion information on new e-books. Even if the user is not an author's follower, the author can still reach the user by sending them an email or promotional message.

Table 7 provides algorithm performance statistics with the test dataset for the author network, and Table 8 provides the same information for the reader network. The average accuracy was 71.6% across all algorithms in the author network, ranging from 68% to 75%. The model generated by the SVM algorithm showed the highest accuracy (75%), the essential criterion for evaluating performance in this research. The AdaBoost model demonstrated the best recall performance, outperforming the 58% average by classifying authors accurately 73% of the time.

For the reader network, although SVM also had the best performance in accuracy and precision, all methods were not significantly different in their performance for accurately predicting e-book purchases. Furthermore, all methods performed poorly with an average 60% accuracy ranging from 58.23% to 63.03%, 11% below the author network. GA-ensemble attained the best recall performance classifying around 20.1% accurately in the test set, on average, 52% below the author network.

Table 7 Author (equivalent) network performance measures. Percentage attained and algorithm success ranking in parenthesis

Learning technique	*Accuracy*	*Recall*	*Accuracy + recall ranking*	*Precision*	*F-measure*	*Overall ranking*
AdaBoost	72.73 (3)	87.58 (1)	1	53.60 (5)	66.50 (1)	1
Decision tree	69.91 (5)	38.03 (6)	6	96.84 (2)	54.91 (5)	6
GA-ensemble	73.93 (2)	65.36 (3)	2	56.82 (4)	60.79 (2)	2
Neural network	67.88 (6)	72.55 (2)	4	48.68 (6)	58.27 (3)	5
Random forest	70.37 (4)	38.98 (5)	5	100.00 (1)	55.10 (4)	3
SVM	74.75 (1)	48.37 (4)	2	61.67 (3)	54.21 (6)	3
Average	*71.60*	*58.48*		*69.60*	*58.30*	

The difference in prediction accuracy between the networks suggests that purchase behaviour prediction in an SNS is significantly affected by the users' social structure. Further, the author network's higher prediction accuracy by any of the chosen algorithms supports hypothesis three – social influence is more significant in social equivalent or homogeneous networks than socially heterogeneous networks.

Table 8 Reader (inequivalent) network performance measures. Percentage attained and algorithm success ranking in parenthesis

Learning technique	*Accuracy*	*Recall*	*Accuracy + recall ranking*	*Precision*	*F-measure*	*Overall rank*
AdaBoost	60.03 (4)	1.17 (5)	5	29.55 (5)	2.25 (5)	5
Decision tree	60.87 (3)	1.44 (4)	4	48.15 (3)	2.80 (4)	4
GA-ensemble	59.46 (5)	20.11 (1)	3	46.47 (4)	28.07 (1)	3
Neural network	58.23 (6)	1.11 (6)	6	12.20 (6)	2.03 (6)	6
Random forest	61.08 (2)	7.42 (2)	1	58.26 (2)	13.16 (2)	1
SVM	63.03 (1)	5.43 (3)	1	100.00 (1)	10.29 (3)	1
Average	*60.45*	*6.14*		*40.11*	*9.77*	

To examine the relationship further, hypothesis testing under the alternate hypothesis author (equivalent) networks have more social links than reader (inequivalent) networks, utilised the t-test. The test statistics (p-values) for degree and weighted degree measures are 35.01 (< 0.001) and 35.00 (< 0.001), respectively. The results support hypothesis three – more social links and interactions exist among structural equivalent entities than otherwise.

5.1 Additional analysis

This research also examined the social influence effects to corroborate those results found in the previous section. This analysis used logistic regression with network measures and subnetwork types as independent variables. Membership was transformed into binary coding where 1 = author and 0 = reader. The logistic regression result suggests that the author's network had a more significant effect on purchases (Beta: 2.595, p-value: < 0.001, odds ratio: 13.403). This outcome further supports hypothesis three – socially equivalent people have more significant social influence and facilitate more accurate predictions.

We also examined the social equivalence effect by comparing the prediction accuracy between the equivalent and inequivalent groups (see Table 9). This analysis utilises the same predictive variables as Tables 7 and 8; however, the data are not split into socially equivalent networks. Instead, the entire dataset was split into modelling and testing subsets, each with two years of data. Then algorithm performance measures were collected and compared to the social equivalent subgroups.

The mixed network obtains significantly higher average accuracy (68.8%) over the inequivalent network (60.5%) (see Tables 9 and 8). However, lower accuracy occurs in the mixed network versus the equivalent network (68.8% mixed versus 71.6% equivalent) (see Tables 9 and 7). This finding supports hypothesis four; social equivalence status biases the prediction accuracy. This discovery is particularly troubling because the mixed dataset contains 24% more inequivalent (reader) than equivalent (author) observations. Our results suggest that predicting purchase behaviours in an SNS without considering users' social structure leads to a biased interpretation.

Table 9 Network performance measures for all users - without social equivalence

Learning technique	*Accuracy*	*Recall*	*Precision*	*F-measure*
AdaBoost	70.63	58.49	71.49	64.34
Decision tree	72.94	64.48	72.69	68.34
GA-ensemble	65.29	59.62	62.18	60.88
Neural network	58.45	10.30	41.67	16.52
Random forest	70.91	71.12	60.58	65.40
SVM	74.41	95.57	64.74	77.19
Average	*68.77*	*59.93*	*62.23*	*58.78*

6 Discussion

The results hold several implications. First, structurally equivalent people on an SNS (i.e., author network) have more social links, supporting the structural equivalence model. In comparison, an entity in the author network had about 50 connections on average, while the reader network had 11. This finding suggests that social position is a source of homogeneity or similarity among social entities in a network. The similarity offers a foundation for establishing social links. Social influence by structural equivalence is essential for explaining or predicting social entities' economic behaviour. Those who are structurally equivalent have more social influence over others due to their similarity in product preference, attitude, or belief even when their social interactions are weak.

These results suggest that predicting a social entities' behaviour without considering social links or interactions could easily lead to a biased interpretation. For example, when data is mixed with homogeneous and heterogeneous groups, the predictive power comes from the homogeneous group. In contrast, the heterogeneous group generates a kind of noise. Thus, examining or predicting a social entity's behaviour based on social interactions without considering structural equivalence may exaggerate the social influence effect because of the social interactions by structural inequivalent entities. As measured through social links or interactions, heterogeneous social entities' product preference or attitude is not proportional to their social influence. Not considering social equivalence could mislead prediction analysis by either underestimating the social influence effect exerted by social equivalent groups or overestimating socially inequivalent groups.

In contrast, although they show infrequent social interaction, homogenous social entities could have substantial social influence. Previous studies have often treated social entities as homogenous groups and analysed social interactions for predicting their behaviours. However, structural equivalence could be a better social influence predictor. In this light, effectively segmenting social positions in a network and comparing social influences by positions is a crucial future research topic.

Last but not least, all algorithm performance was relatively stable and consistent in the author network across all four measurement statistics. By ranking them from 1-to-6 (Tables 7 and 8), the resulting rank structure holds more or less constant under any statistic from accuracy to the F-measure. However, all algorithms were far less accurate in the reader network, as rated by recall and the F-measure. SVMs 100% performance, as measured by precision, is combined with the best performance as measured by accuracy; however, the algorithm has comparatively poor performances as measured by recall (5%) and modest F-measures (10%), suggesting the precision number lacks reliability. Further, given the reader network size in the test set (n > 17k), 100% precision, while technically correct and desirable, seems unobtainable, particularly given the moderate recall scores.

Similarly, in the reader network (Table 8), GA-ensemble turned in the second-to-last performance measured by accuracy only to turn in the best performance measured by recall and the F-measure. It is as if the measurements stand opposed with accuracy and precision telling one story, while recall and F-measure tell a different story. It is assumed that the reader network performance characteristics are derived from the low connectivity observed in the network. The low connectivity gives the dataset a diverging nature that approaches random noise. The network's nature warrants further study; however, that is left to future work.

These results suggest that algorithm selection needs to carefully consider the analysis objective and the social structures nature in which the predictions will occur. AdaBoost and GA-ensemble both had satisfactory performance under any measurement criteria in the socially equivalent network; however, SVM and random forest performed best in the inequivalent network.

7 Theoretical and practical implications

7.1 Theoretical implications

This study has several implications for social network studies. First, we suggest an alternative way to predict product purchases in a social network by considering transactions in a network as nodes and applying structural measurements. This approach is believed to be effective, especially when accessing data on social relationships is limited. Second, our study exhibits the importance of examining users' social structural equivalence in predicting their behaviour in a social network. Previous studies often predict users' behaviour by analysing their interactions in interaction frequency terms, the interaction direction, or the information amount. However, such an interaction centred approach can mislead because the interactions are dependent upon their social structure. It is essential to analyse or control users' social structural equivalence while predicting behaviour. Finally, our study also demonstrates the importance of occupation in determining social equivalence. Authors act as a homogeneous group and show more community (they contribute to the network), which leads to frequent interactions. On the other hand, readers are a more diverse, heterogeneous group, partly due to their weak similarity (they consume). Hence, future studies need to identify important factors that help determine the users' social equivalence in the network.

7.2 Practical implications

This study reveals several practical implications. First, when building algorithms to forecast sales performance, sellers on e-commerce platforms should consider the extant measures such as customers' past purchases and their product preferences and network metrics. Including the network features in training the algorithms would help improve the prediction accuracy by a large margin. A more precise sales forecasting could help the sellers develop more effective business strategies.

Second, sellers may segment customers based on their social equivalence. More specifically, they should take the buyers' role on the network into account. SNS users who have more homogenous characteristics may have more frequent interactions so that they are more likely to purchase the products within their circle. For example, an indie moviemaker may be more likely to consume a movie produced by another indie moviemaker. Conducting market segmentation based on this user characteristic may help the sellers more accurately target marketing.

Last but not least, SNS platform designers and data scientists should consider adjusting the algorithms (e.g., GA-ensemble versus SVM) in different situations (i.e., socially equivalent network versus socially inequivalent network) to maximise the predictive performance. Albeit technical, such algorithmic applications should be a common practice in a company that has a well-trained IT department, especially for an SNS platform.

8 Conclusions

This study examined social influence by structural equivalence on e-book purchasing forecasts. Social influence is more significant in social equivalent or homogeneous than heterogeneous networks. Socially equivalent people have more significant social influence and facilitate more accurate predictions. Further, failure to consider social equivalence leads to biased predictions.

The study has several limitations. First, although social relationship estimates from product purchases assumed that the purchase reflects social interaction, the assumption was not examined. Second, since users in the data made limited e-books purchases, the estimation of the structural measure could be biased. Finally, the analysis only considered book transactions in the novel genre; therefore, the results cannot be generalised to all book types. Finally, where the analysis only considered e-book transactions on a particular SNS, generalising the results is left to future research. That research needs to examine e-books across genres, across various platforms, and then begin exploring asset sales other than e-books.

References

Almgren, K. and Lee, J. (2016) 'An empirical comparison of influence measurements for social network analysis', *Social Network Analysis and Mining*, Vol. 6, No. 1, pp.1–18.

Anderson, P. (2016) 'Bowker, now cites at least 625,327 US indie books published in 2015', *Publishing Perspectives* [online] http://publishingperspectives.com/2016/10/bowker-indie-titles-2015-isbn/#.WGaOf_krKUk (accessed 10 July 2019).

Aral, S. and Walker, D. (2011) 'Creating social contagion through viral product design: a randomised trial of peer influence in networks', *Management Science*, Vol. 57, No. 9, pp.1623–1639.

Bhatt, R., Chaoji, V. and Parekh, R. (2010) 'Predicting product adoption in large-scale social networks', in *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*, pp.1039–1048.

Bookbaby (2017) Self-publishing survey results [online] https://www.bookbaby.com/pages/pdf/official-self-publishing-survey-results.pdf (accessed 15 August 2019).

Breiman, L. (2001) 'Random forests', *Machine Learning*, Vol. 45, No. 1, pp.5–32.

Breiman, L., Friedman, J., Olshen, R. and Stone, C. (1984) *Classification and Regression Trees*, Belmont, Wadsworth, California.

Burkhardt, M.E. (1994) 'Social interaction effects following a technological change: a longitudinal investigation', *Academy of Management Journal*, Vol. 37, No. 4, pp.869–898.

Burt, R.S. (1982) *Toward a structural theory of action Network: Models of Social Structure, Perception and Action*, Academic Press, New York.

Burt, R.S. (1987) 'Social contagion and innovation, cohesion versus structural equivalence', *American Journal of Sociology*, Vol. 92, No. 6, pp.1287–1335.

Capozzi, F., Becchio, C., Willemse, C. and Bayliss, A.P. (2016) 'Followers are, not followed: observed group interactions modulate subsequent social attention', *Journal of Experimental Psychology: General*, Vol. 145, No. 5, pp.531–535.

Cha, M., Haddadi, H., Benevenuto, F. and Gummadi, K.P. (2010) 'Measuring user influence in twitter: The million follower fallacy', in *Proceedings of the 4th International AAAI Conference on Weblogs and Social Media*, pp.10–17.

Chang, X. and Li, J. (2019) 'Business performance prediction in location-based social commerce', *Expert Systems with Applications*, Vol. 126, pp.112–123.

Chen, J. and Shen, X.L. (2015) 'Consumers' decisions in social commerce context: an empirical investigation', *Decision Support Systems*, Vol. 79, pp.55–64.

Chiu, C.M., Cheng, H.L., Huang, H.Y. and Chen, C.F. (2013) 'Exploring individuals' subjective well-being and loyalty towards social network sites from the perspective of network externalities: the Facebook case', *International Journal of Information Management*, Vol. 33, No. 3, pp.539–552.

Chu, S.C. and Kim, Y. (2011) 'Determinants of consumer engagement in electronic word-of-mouth (eWOM) in social networking sites', *International Journal of Advertising*, Vol. 30, No. 1, pp.47–75.

Chung, Y., Li, Y. and Jia, J. (2021) 'Exploring embeddedness, centrality and social influence on backer behaviour: the role of backer networks in crowdfunding', *Journal of the Academy of Marketing Science*, pp.1–22.

Cortes, C. and Vapnik, V. (1995) 'Support-vector networks', *Machine Learning*, Vol. 20, No. 3, pp.273–297.

Dabbous, A., Aoun Barakat, K. and Merhej Sayegh, M. (2020) 'Social commerce success: Antecedents of purchase intention and the mediating role of trust', *Journal of Internet Commerce*, Vol. 19, No. 3, pp.262–297.

Dhar, V., Geva, T., Oestreicher-Singer, G. and Sundararajan, A. (2014) 'Prediction in economic networks', *Information Systems Research*, Vol. 25, No. 2, pp.264–284.

Fang, X., Hu, P.J.H., Li, Z. and Tsai, W. (2013) 'Predicting adoption probabilities in social networks', *Information Systems Research*, Vol. 24, No. 1, pp.128–145.

Fenzl, T. and Pelzmann, L. (2012) 'Psychological and social forces behind aggregate financial market behaviour', *Journal of Behavioural Finance*, Vol. 13, No. 1, pp.56–65.

Freeman, L.C. (1978) 'Centrality in social networks conceptual clarification', *Social Networks*, Vol. 1, No. 3, pp.215–239.

Freund, Y. and Schapire, R. (1997) 'A decision-theoretic generalisation of online learning and an application to boosting', *Journal of Computer and System Sciences*, Vol. 55, No. 1, pp.119–139.

Friedkin, N.E. (1984) 'Structural cohesion and equivalence explanations of social homogeneity', *Sociological Methods & Research*, Vol. 12, No. 3, pp.235–261.

Friedkin, N.E. (1993) 'Structural bases of interpersonal influence in groups: a longitudinal case study', *American Sociological Review*, Vol. 58, No. 6, pp.861–872.

Gibreel, O., AlOtaibi, D.A. and Altmann, J. (2018) 'Social commerce development in emerging markets', *Electronic Commerce Research and Applications*, Vol. 27, No. 4, pp.152–162.

Goh, T.T., Yang, B., Dai, X. and Jin, D. (2017) 'A study of purchase influence and behavioural intention on the adoption of electronic word of mouth (eWOM) systems', *Journal of Electronic Commerce in Organisations*, Vol. 15, No. 3, pp.14–32.

Gunawan, D.D. and Huarng, K.H. (2015) 'Viral effects of social network and media on consumers' purchase intention', *Journal of Business Research*, Vol. 68, No. 11, pp.2237–2241.

Guo, S., Wang, M. and Leskovec, J. (2011) 'The role of social networks in online shopping: information passing, price of trust and consumer choice', in *Proceedings of the 12th ACM Conference on Electronic Commerce*, pp.157–166.

Hajli, M.N. (2014) 'A study of the impact of social media on consumers', *International Journal of Market Research*, Vol. 56, No. 3, pp.387–404.

Hajli, N., Sims, J., Zadeh, A.H. and Richard, M.O. (2017) 'A social commerce investigation of the role of trust in a social networking site on purchase intentions', *Journal of Business Research*, Vol. 71, pp.133–141.

Han, S. and Park, K.G. (2020) 'Social network services and their effects on network social capital: an instrumental variables approach', *International Journal of Mobile Communications*, Vol. 18, No. 4, pp.386–404.

Handarkho, Y.D. (2020) 'Impact of social experience on customer purchase decision in the social commerce context' *Journal of Systems and Information Technology*, Vol. 22, No. 1, pp.47–71.

Harkola, J. and Greve, A. (1995) 'Diffusion of technology: cohesion or structural equivalence?', in *Academy of Management Proceedings*, pp.422–426.

Hinz, O., Schulze, C. and Takac, C. (2014) 'New product adoption in social networks: why direction matters', *Journal of Business Research*, Vol. 67, No. 1, pp.2836–2844.

Hinz, O., Skiera, B., Barrot, C. and Becker, J.U. (2011) 'Seeding strategies for viral marketing: an empirical comparison', *Journal of Marketing*, Vol. 75, No. 6, pp.55–71.

Hollenbaugh, E.E. and Ferris, A.L. (2014) 'Facebook self-disclosure: examining the role of traits, social cohesion and motives', *Computers in Human Behaviour*, Vol. 30, pp.50–58.

Hromic, H. and Hayes, C. (2019) 'Characterising and evaluating dynamic online communities from live microblogging user interactions', *Social Network Analysis and Mining*, Vol. 9, No. 1, p.30.

Hu, X., Chen, X. and Davison, R.M. (2019) 'Social support, source credibility, social influence and impulsive purchase behaviour in social commerce', *International Journal of Electronic Commerce*, Vol. 23, No. 3, pp.297–327.

Huang, Z. and Benyoucef, M. (2013) 'From e-commerce to social commerce: a close look at design features', *Electronic Commerce Research and Applications*, Vol. 12, No. 4, pp.246–259.

Hudson, S., Huang, L., Roth, M.S. and Madden, T.J. (2016) 'The influence of social media interactions on consumer–brand relationships: a three-country study of brand perceptions and marketing behaviours', *International Journal of Research in Marketing*, Vol. 33, No. 1, pp.27–41.

Katona, Z., Zubcsek, P.P. and Sarvary, M. (2011) 'Network effects and personal influences: the diffusion of an online social network', *Journal of Marketing Research*, Vol. 48, No. 3, pp.425–443.

Kietzmann, J.H., Hermkens, K., McCarthy, I.P. and Silvestre, B.S. (2011) 'Social media? Get serious! Understanding the functional building blocks of social media', *Business Horizons*, Vol. 54, No. 3, pp.241–251.

Kim, Y.A. and Srivastava, J. (2007) 'Impact of social influence in e-commerce decision making', in *Proceedings of the ninth international conference on Electronic commerce*, pp.293–302.

Kiss, C. and Bichler, M. (2008) 'Identification of influencers—measuring influence in customer networks', *Decision Support Systems*, Vol. 46, No. 1, pp.233–253.

Krause, J., Croft, D.P. and James, R. (2007) 'Social network theory in the behavioural sciences: potential applications', *Behavioural Ecology and Sociobiology*, Vol. 62, No. 1, pp.15–27.

Kumar, R., Novak, J. and Tomkins, A. (2010) 'Structure and evolution of online social networks', in *Link Mining: Models, Algorithms and Applications*, pp.337–357, Springer, New York.

Lee, S. and Choeh, J.Y. (2020) 'Using the social influence of electronic word-of-mouth for predicting product sales: the moderating effect of review or reviewer helpfulness and product type', *Sustainability*, Vol. 12, No. 19, p.7952.

Leeraphong, A. and Mardjo, A. (2013) 'Trust and risk in purchase intention through online social network: a focus group study of Facebook in Thailand', *Journal of Economics, Business and Management*, Vol. 1, No. 4, pp.314–318.

Li, Y.M. and Lai, C.Y. (2014) 'A social appraisal mechanism for online purchase decision support in the micro-blogosphere', *Decision Support Systems*, Vol. 59, pp.190–205.

Liang, T.P., Ho, Y.T., Li, Y.W. and Turban, E. (2011) 'What drives social commerce: the role of social support and relationship quality', *International Journal of Electronic Commerce*, Vol. 16, No. 2, pp.69–90.

Lin, J., Luo, Z., Cheng, X. and Li, L. (2019a) 'Understanding the interplay of social commerce affordances and swift guanxi: An empirical study', *Information & Management*, Vol. 56, No. 2, pp.213–224.

Lin, K.Y. and Lu, H.P. (2011) 'Why people use social networking sites: an empirical study integrating network externalities and motivation theory', *Computers in human behaviour*, Vol. 27, No. 3, pp.1152-1161.

Lin, X., Wang, X. and Hajli, N. (2019b) 'Building e-commerce satisfaction and boosting sales: the role of social commerce trust and its antecedents', *International Journal of Electronic Commerce*, Vol. 23, No. 3, pp.328–363.

Lipsman, A. (2019, June 27) Global ecommerce 2019: Ecommerce continues strong gains amid global economic uncertainty', *eMarketer* [online] https://www.emarketer.com/content/global-ecommerce-2019 (accessed 10 July 2019).

Liu, H., Chu, H., Huang, Q. and Chen, X. (2016) 'Enhancing the flow experience of consumers in China through interpersonal interaction in social commerce', *Computers in Human Behaviour*, Vol. 58, pp.6–314.

Lu, B., Fan, W. and Zhou, M. (2016) 'Social presence, trust and social commerce purchase intention: an empirical research', *Computers in Human Behaviour*, Vol. 56, pp.225–237.

Ma, L., Krishnan, R. and Montgomery, A. (2010) *Homophily or Influence? An Empirical Analysis of Purchase within a Social Network*, Working paper, Carnegie Mellon University.

McCulloch, W. and Walter, P. (1943) 'A logical calculus of ideas immanent in nervous activity', *Bulletin of Mathematical Biophysics*, Vol. 5, No. 4, pp.115–133.

Molden, D.C. and Dweck, C.S. (2006) 'Finding 'meaning' in psychology: a lay theories approach to self-regulation, social perception and social development', *American Psychologist*, Vol. 61, No. 3, pp.192–203.

Oh, D.Y. and Gray, J.B. (2013) 'GA-Ensemble: a genetic algorithm for robust ensembles', *Computational Statistics*, Vol. 28, No. 5, pp.2333–2347.

Pallotti, F. and Lomi, A. (2011) 'Network influence and organisational performance: the effects of tie strength and structural equivalence', *European Management Journal*, Vol. 29, No. 5, pp.389–403.

Peng, S., Wang, G. and Xie, D. (2016) 'Social influence analysis in social networking big data: opportunities and challenges', *IEEE network*, Vol. 31, No. 1, pp.11–17.

Peng, S., Yang, A., Cao, L., Yu, S. and Xie, D. (2017) 'Social influence modeling using information theory in mobile social networks', *Information Sciences*, Vol. 379, pp.146–159.

Rapp, A., Beitelspacher, L.S., Grewal, D. and Hughes, D.E. (2013) 'Understanding social media effects across seller, retailer, and consumer interactions', *Journal of the Academy of Marketing Science*, Vol. 41, No. 5, pp.547–566.

Raven, B.H. (1964) *Social Influence and Power*, California University, Los Angeles.

See-To, E.W. and Ho, K.K. (2014) 'Value co-creation and purchase intention in social network sites: the role of electronic word-of-mouth and trust–a theoretical analysis', *Computers in Human Behaviour,* Vol. 31, pp.182–189.

Statista (2019) 'eBooks worldwide', *Statistia Statistics* https://www.statista.com/outlook/213/100/ebooks/worldwide (accessed 15 August 2019).

Stephen, A.T. and Toubia, O. (2010) 'Deriving value from social commerce networks', *Journal of Marketing Research*, Vol. 47, No. 2, pp.215–228.

Teh, P.L., Ahmed, P.K. and Tayi, G.K. (2015) 'Generation-Y shopping: the impact of network externalities and trust on adoption of social commerce', *International Journal of Electronic Business,* Vol. 12, No. 2, pp.117–141.

Trusov, M., Bucklin, R.E. and Pauwels, K. (2009) 'Effects of word-of-mouth versus traditional marketing: Findings from an internet social networking site', *Journal of Marketing*, Vol. 73, No. 5, pp.90–102.

US Small Business Administration (2018) 2018 Small business profile, [online] https://www.sba.gov/sites/default/files/advocacy/2018-Small-Business-Profiles-US.pdf (accessed 15 August 2019).

Varian, H.L. (2005, August 25) 'Technology levels the business playing field', *New Your Times*, [online] https://www.nytimes.com/ (accessed 14 September 2019).

Verbraken, T., Goethals, F., Verbeke, W. and Baesens, B. (2014) 'Predicting online channel acceptance with social network data', *Decision Support Systems*, Vol. 63, pp.104–114.

Wang, X., Yu, C. and Wei, Y. (2012) 'Social media peer communication and impacts on purchase intentions: A consumer socialization framework', *Journal of Interactive Marketing*, Vol. 26, No. 4, pp.198–208.

Wang, Y. and Yu, C. (2017) 'Social interaction-based consumer decision-making model in social commerce: The role of word of mouth and observational learning', *International Journal of Information Management,* Vol. 37, No. 3, pp.179–189.

Worldbank (2021) Small and medium enterprises (SMEs) finance: improving SMEs' access to finance and finding innovative solutions to unlock sources of capital [online] https://www.worldbank.org/en/topic/smefinance (accessed 25 February 2020).

Xiao, L., Fu, B. and Liu, W. (2018) 'Understanding consumer repurchase intention on O2O platforms: an integrated model of network externalities and trust transfer theory', *Service Business*, Vol. 12, No. 4, pp.731–756.

Yoo, Y. and Alavi, M. (2001) 'Media and group cohesion: relative influences on social presence, task participation and group consensus', *MIS Quarterly*, Vol. 25, No. 3, pp.371–390.

Young, J. (2021) 'Global online sales reach nearly $4.29 trillion in 2020', *Digital Commerce 360*, [online] https://www.digitalcommerce360.com/article/global-ecommerce-sales/#:~:text=Consumers%20worldwide%20spent%20nearly%20%244.29,to%20Digital%20Commerce%20360%20estimates (accessed 30 May 2021).

Zhou, T. (2019) 'The effect of social interaction on users' social commerce intention', *International Journal of Mobile Communications*, Vol. 17, No. 4, pp.391–408.

Zhu, Y.Q., Amelina, D. and Yen, D.C. (2020) 'Celebrity endorsement and impulsive buying intentions in social commerce-the case of Instagram in Indonesia: celebrity endorsement', *Journal of Electronic Commerce in Organisations*, Vol. 18, No. 1, pp.1–17.