# Use of the Pathfinder Network scaling to measure online customer reviews: A Theme Park Study

To cite this paper: Qun Ren, Feifei Xu, Xiaoyi Ji. 2019 Use of the pathfinder network scaling to measure online customer reviews: a them park study. Strategic Change. DOI: 10.1002/jsc.2288

One-sentence summary: We use pathfinder network scaling approach to measure and evaluate Theme Park visitors’ online reviews.

# key Points:

Pathfinder Network Scaling (PFNETs) as an effective tool of big data analytics can be used to identify unobserved meaningful interrelationships between concepts. Although there are many research analysing online reviews, this study is the first attempt to use an analytical approach of Pathfinder Network Scaling (PFNETs) to explore online reviews in Theme park visitors experiences. The paper collects 14,142 effective reviews of the World’s first Disneyland in California from Trip Advisor. Using parallel and similarity comparison in Pathfinder scaling, four individually but fully connected networks were generated to reveal different visitors’ experiences in different segments. The findings indicate the dissimilarity of concept relatedness between different segments and revealed the knowledge gap of marketing to different segments in theme parks.

As the increasing growth and wide use of social network sites (hereinafter, SNS) has become a global phenomenon, visitors’ online reviews (or digital texts) on the social network sites (hereinafter, SNS) has a ubiquitous influence on a potential visitor’s understanding to a Theme Park’s service quality, relationship with customers as well as making his/her purchase decision (Arenas-Gaitan et al., 2013; Wang and Chang, 2013). It is by no means clear, more and more organizations in the hospitality industry have regarded online reviews as the key source of knowledge (Roblek et al.,2013). Consequently, there is a growing research interest on analysing online reviews to understand customers’ needs, to build brand loyalty and to encourage customer engagement (Sánchez-Franco, et al 2017; Crosby et al., 1990; Hennig-Thurau and Klee, 1997).

Tourism has been one of the areas that online reviews have been explored extensively. For example, Xiang et. al. (2015) used reviews from Trip Adsivor to study hotel staying experiences. Geetha et al. (2017) used reviews to explore customers sentiment with hotels. However, previous research on theme park experiences have mainly been relying on traditional surveys and interviews (See Pikkemaat and Schuckert 2007; Geissler and Ruck 2011; Chiappa et al. 2013), therefore, inevitably have their limitations from traditional research methods. Therefore, it would be helpful to evaluate and gain insight from large amount of unstructured online reviews from consumers (Asmith, et al, 2017). Recognizing this research gap, the paper hopes to contribute to the studies on customer experiences in theme parks by analysing online reviews.

In the field of marketing, over the years, researchers have become increasingly aware of the significance of understanding customer segments (Rau et al, 2008). Segmentation plays an important role in marketing as different groups of customers will have different needs, motivation and behaviour, therefore, shaping different experiences (Xu et al. 2011). How to identify these experiences that can be shared by a certain group becomes critical to marketing. With the advent of Web 2.0, using a large set of user generated content makes it possible in identifying unstructured hidden patterns when exploring customer experiences.

Among many methods of big data analytical tools, pathfinder scaling is very effective. It is used to set a fully connected network during the data analysis, which indicates the importance of these terms and their relationship in the perception of different customer segments. Pathfinder Network Scaling (PENETS) was originally developed in the domain of psychology. It was well established and widely used to evaluate expertise difference between groups of individuals, for instance, teachers and students (Housner, et al., 1993; Goldsmith et al., 1990). Pathfinder nets are network graphs representing semantic distances between concepts by calculating and mapping relatedness (Gonzalvo et al., 1994; Jonassen, 1993). PFNETS presents a desirable solution for analysing online reviews as can help automatically infer the knowledge structure from large amounts of unstructured online content authored by theme park visitors and shown on TripAdvisor SNS; In addition, it can be an effective tool to use the minimum spanning tree structure to derive and reveal cognitive differences between the identified segments. Although previous studies have applied different methods to analyse online customer reviews in the areas of hospitality and tourism (Sánchez-Franco et al, 2016; Min et al, 2015; Ho-Dac, 2014; Felbermayr & Nanopoulos,2016), to the best of our knowledge to date, no previous studies have been found using pathfinder scaling in tourism studies.

As the application of ubiquitous Internet and the acceleration of mobile internet usage continue to increase globally, there is an ever-increasing power switch between theme park and its visitors. Previously, visitors are largely dependent on the information source to get what they wanted, now digital visitors have their capability to get to know and share their information online to search for differentiated offerings to cater to their individuality.

Therefore, segmentation of visitors is widely accepted and adopted by theme parks. Central to this concern is how theme parks can segment their market with the big data to work with and profile the customers accurately and comprehensively (Frochot & Morrison, 2000; Prayag & Hosany, 2014; Prayag, Disegna, Cohen, & Yan, 2015). The key is that theme parks need to respond and reflect the growing desire for authentic theme park visit, differentiate and personalize their services so as to provide relevant and unique services/product to the diversified target markets.

Smith (1956) articulated the necessity of market segment redefinition since the defined market segments will keep on changing over different periods of boom and crisis. (Freytag Clarke, 2001). The set of product mix and marketing mix will change following the business environment dynamics and customers’ need changes. Hence, the key to winning and retaining loyalty and trust is to place match the market offering with the market segments' changing expectations under the big context of digital technology ubiquity.

The research is developed within the context of theme parks. As the world’s first theme park, Disneyland in California attracts 17.94 million visitors a year (TEA/AEOM, 2016). However, there are few studies exploring online reviews of Disneyland. Therefore, the aim of this paper is to use pathfinder scaling to understand customer experiences in different segments of theme park visitors. The following research questions will be explored in this research: 1) to identify the customer experiences in different segments; 2) to compare underlining differences between segments using parallel comparison and similarity comparison.

## **Literature Review:**

*Customer Experiences*

Customer experience is a critical part in services. It mainly includes three aspects: the natural environment, the interaction with employees and the interaction of other customers (Tasci & Mailman, 2019). Although different researchers include different elements in natural environment, it mainly include physical facilities and equipment and staff (Parasuraman et al. 1988). Bitner (1992) refers this as "servicescape". When exploring the essence of service quality and customer satisfaction, many scholars stressed the significance of considering intangibility, heterogeneity, and inseparability, i.e. the three core service characteristics of service quality. (Parasuraman et al., 1988; Tsang et al., 2012). Therefore, a lot of studies have focused on the importance of customer staff interaction throughout the service process suggesting that customers would often evaluate the experience based on staff’s attitude, performance and friendliness (Homburg, Koschate and Hoyer, 2005). However, whether different segments perceive the same service differently, in other words, will different segments pay attention to different aspects of customer experiences is not clear.

Tourism is what Pine and Gillmore (1999) referred as experience industry. However, the industry has been previously criticised as focusing on promoting the physical attributes of the destination while ignoring the individual experiences, fulfilment and rejuvenation (King, 2002). Williams (2006) also urge attention needs to be paid on experiential marketing. Effective experiential marketing needs to be based on research into the consumer experience. Drawing on Kapferer’s (1998) prism of brand identity, Morgan and Xu (2009) suggest tourist experiences may include physical attributes, destination image, social interactions, cultural interactions, benefits and meanings. Physical attributes and destination image are extrinsic factors similar as Echtner and Ritchie’s destination image model (2003). While benefits and meanings are inward driven factors which emphasis on emotional or psychological benefits derived from the experience (Morgan and Xu, 2009). In between, the interaction usually includes social and cultural interactions with the destination and with the travel companion. In conclude, the tourist experiences can be included as outward driven (attribute based), inward driven(emotional based) and interaction(social cultural interaction) driven. However, limited research has implied this with different segments (Xu et al. 2011), whether a particular group will pay attention to any certain aspect of inward, outward or interaction driven factors in unclear.

*Strategic change in theme parks and Disney*

Theme parks play a key role in tourism industry as they contribute significantly to global revenue generation (Sievänen et al., 2011). Successful interplay between a theme park’s competitive strategies and positive travel experience will lead to high tourist satisfaction and value creation (Rashidirad et al., 2017). With the popularity of theme parks, visitor experiences in theme parks became a research focus for marketing. In the extremely competitive theme park tourism market, it is increasingly important for theme park managers to gain a good understanding of what constitutes a satisfactory service experience can from customers’ point of view (Birenboim et al., 2013), re-evaluate what their offerings appeal to customers and hence identify the gap in the market (Crick and Crick, 2016).

Recently, the growing ubiquity of high-speed mobile Internet drives the Disneyland’s strategic changes. In terms of this, Disneyland is transforming quickly to adapt and continue move online. This can be exemplified using an official app for the Disneyland resort. This specific app can be downloaded for either Android or iPhone, which enables the face-to-face processes to be replaced by digital alternatives. This app not only can tell you the most accurate wait times for rides, it also has the digitally-together capability. That is to say, it can help you to skip the ticket lines by buying park tickets online make dinning reservations, see restaurant menu, Check Disney FASTPASS Return Times. All these mobile experiences will be enchanting and enhance visitors’ retention.

The financial crash in 2008 reminded consumers and has shifted vast consumers’ favour from luxurious products and services to simplified level of status which indicate authenticity and individuality. This growing desire for authentic travel is mirrored as Uber and Airbnb are thriving businesses around Disneyland. More and more visitors are used to favour simplicity and individuality by creating their own accommodation package as an alternative to Disneyland Hotels by integrating: 1) convenient and value for money service from Uber and Airbnb and 2) relatively low -cost and authentic hotels which around 2 or 3 miles from Disneyland.

Previous studies on visitor experiences in theme parks indicate that, following Two-dimensional matrix IPA model (Martilla and James, 1977) and SERVQUAL instrument (Zeithaml et al., 1988) which attempted to set five variables for evaluating a service (i.e. tangibility, reliability, assurance, responsiveness and empathy), many other scholars kept on developing refined tools and adding new variables in order to measure and evaluate a service quality in a more comprehensive approach in the theme park industry (Brown et al.1993; Oh, 2001; Van Ree, 2009; Slak Valek et al., 2014) while other researchers attempted to investigate customer’s satisfaction from the perspective of safety, security and environment uniqueness (Pikkemaat and Schuckert, 2007). Similarly, Tsang et al (2012) notified that assurance, empathy and responsiveness are key factors to predict the theme park visitors’ satisfaction levels. In addition, Factor analysis and cluster analysis is used by some scholars for measuring different variables (such as time, cost) with regards to theme park’s segmentation analysis (Vassiliadis et al. ,2013; Fotiadis et al,2017).

However, previous research on theme park experiences have mainly used traditional surveys and interviews (See Pikkemaat and Schuckert 2007; Geissler and Ruck 2011; Chiappa et al. 2013), therefore, inevitably have their limitations from traditional research methods. With the fast development of big data analytics, particularly, Web 2.0 has made user generated contents (UGC) available, it is possible to use large sets of UGC in identifying unstructured hidden patterns. Although previous studies have applied different methods to analyse online customer reviews in the areas of hospitality and tourism (Sánchez-Franco et al, 2016; Min et al, 2015; Ho-Dac, 2014; Felbermayr & Nanopoulos, 2016), so far, limited efforts has been made regarding how to help managers in the theme parks industry to understand online customer experiences in different segments with a network analysis approach to identity concepts and mind maps.

*Pathfinder scaling approach*

Pathfinder Network Scaling (PFNETS) was first developed in the domain of psychology. It was well established and widely used to evaluate expertise difference between groups of individuals, for instance, teachers and studens (Housner, et al., 1993; Goldsmith et al., 1991). Pathfinder nets are network graphs representing semantic distances between concepts by calculating and mapping relatedness (Gonzalvo et al., 1994; Jonassen, 1993). With the use of a shortest path approach, PFNETS transform related ratings into a visible link-weighted networks. Like concept maps, concepts are represented as nodes while relationships between concepts are regarded as links. PFNETs discover significant links by eliminating edges with a distance exceeding that of alternate paths. Based on the assumption that the degree of relatedness indicates the psychological distance; therefore, highly related concepts are separated by fewer links and more links will be between two less related concepts (Day, 2001).

Although Pathfinder scaling techniques have lots of similarity with concept mapping, it was regarded as a better alternative method of concept mapping since much literature recognizes that concept mapping is subjective, time consuming, especially when some educators assessing structural knowledge with students’ use of concept mapping. and its quality maybe be dependent on a lot on students’ understanding levels of concept map. (All, et al,2003; Williams, 2004). In addition, the strength of PENETs lies in its validity to yield structures which can represent and evaluate the organizations of the most frequent terms. Distinguished from structural equation models and other statistical methods, PENETS can effectively convert a set of relatedness judgements into fully connected network which can reflect the structure of conceptual domain better than multidimensional scaling (Goldsmith et al., 1991). It also has higher power in revealing and interpreting the semantic dimensions in contrast with other multidimensional scaling (MDS) representations” (Lau & Yuen, 2009).

PENETS is not only widely accepted and applied in the field of education and psychology, but also well adopted by medical and health professionals in the health domain. For instance, health experts realize that their education material (such as the online vaccine content) should be developed aligning with the consumer’s level of understanding. Therefore, they create online vaccine content with the use of pathﬁnder Network Scaling (PFNETS) for automated text analysis. The resulting graphs generated from application of PFNETs can reveal the evidence of conflict understanding and knowledge gap of vaccine concepts between health consumers and health experts and therefore are able to suggest opportunities to enhance provider-patient communications.

Although well used in other fields, there is very limited number of prior studies in the extant tourism literature on the content and main themes of online user-generated reviews (Sanchez-Franco et al., 2016).

In summary, by review the previous theories, it is obvious that PFNETS could be well adopted by theme parks for analysing online reviews as it can help automatically infer the knowledge structure from large amounts of unstructured online content authored by theme park visitors and shown on TripAdvisor SNS. However, there is a lack of research on the application of PFNETs, especially its application in the theme park industry. Hence, our study below will be based on the data we have collected, in attempting to apply PFNETS as an effective tool to generate the minimum spanning trees so as to derive and reveal cognitive differences between the identified segments.

**Research Method**

### *Data collection: Why using data from TripAdvisor?*

The study collects 14,520 pieces of online reviews from Trip Advisor. Trip Advisor is one of the world’s largest and most popular UGC platform in tourism industry, specializing in travel reviews (Nizmuddin, 2015). In 2017, it published over 500 million tourists’ comments. Therefore, the research uses online reviews from Trip Advisor. From 12 to 22 Aug 2017, a Web crawler software was used to collect online reviews. In total, 14,142 pieces of effective reviews were collected, covering publishing dates range from 1 June 2012 to 30 July 2017. Prior to the data collection, reviews are under the category of different customer segments set by the TripAdvisor system. Under TripAdvisor’s classifies, a database on 5 segments of Disney visitors was established, such as Family (9653), Couple (2599), Solo (315), Business (155) and Friends (1420), however, due to the small number of samples, solo and business were ignored in this study. Therefore, the following study will be based on the three segments, i.e. family, couple and Friends. Each review data contains 1) reviewer ‘name, review date;2) a short title; 3) a rating for their experience; and the entire review texts.

*Data Processing prior to the PFNET’s generation*

The next approach is to process the top 200 TF-IDF terms which are the concepts (i.e. nodes) used for PFNET visualization. Each online review available on the TripAdvisor has been systematically categorized to the segments of: family, couple, friend. Based on these online resources. research starts with adopting text-mining approach. Initially, data were input into a document-term matrix. We remove words that have no useful information with the use of extract list of stop words. The list of stop words is provided on tm packages (cf. R. 3.2.0). Then. We use TF-IDF software to calculate the TF-IDF score (i.e. term frequency (tf) inverse document frequency (idf)). The top TF-IDF scores we use in the rest of the study are the 200 most frequently terms that are cited by the reviewers (See Table 1).

INSERT TABLE 1 NEAR HERE

#### Generation of PFNETs

Overall, in order to make full comparison, based on the tf-idf value, we generated 4 PFNET networks. Each network shows complete inter-connectivity amongst 200 most frequently terms that are cited by the reviewers. Out of the 4 networks generated, one network is PFNET-full which represents the full online review tank; and the other 3 are the visualization of three subsamples: PFNET-Family, PFNET-Couple, PFNET-Friend le (See Figure 1, 2, 3, and 4).

The distance differences of inter-connectivity between all the nodes will reveal importance of the terms in the network and in bridging between other nodes out of the different perception of each customer segment.

### *Analysing the PFNETS*

On analysing the PFNETs, we have applied for parallel comparison and servility comparison respectively. Details can be found in the following Results and Discussion section.

# **Results and Discussion**

Our findings reveal customer experiences shown in PFNET subsamples diverge from the PFNET-total. We clarified our findings based on two types of comparison: Parallel Comparison and Similarity Comparison.

#### Parallel Comparison

Three focal nodes on each PFNET can be found. They are: the highest degree node (the core on each network graph), the median node (a node with smallest maximum distance from other node) and centre node (a node with smallest maximum distance from other node). This research will also analyse the differences based on the customer segments of family, couple and friend and supports that theme park experiences are cognitively structure by using these three different customer segments by using different criteria (shown in Figure 1, Figure 2, Figure 3 and Figure 4).

INSERT FIG 1 NEAR HERE

INSERT FIG 2 NEAR HERE

INSERT FIG 3 NEAR HERE

INSERT FIG 4 NEAR HERE

We conduct the parallel comparison so as to scrutinize and discuss each PFNET with the node discussion: Highest degree of node (i.e. core on the network graph) , Frequency list, centre node (normally the sensory element of reviewers’ experience) and median node. Similarity graph consists of nodes and edges. Each word is a node. In each PFNET, very two nodes are connected via an edge. Each edge has a weight, which is the similarity between the two words. Similarity Graph (which can have circuits) into a minimum spanning tree (which has no circuit) and Similarity table indicates how similar the two words appear in the comments. Centre node refers to the max distance of all terms in Mini Spanning Tree while median node is the sum distance of all terms in Mini spanning tree. Here is the summary of parallel comparison below (See Table 2).

INSERT TABLE 2 NEAR HERE

PFNET-Friend:

The highest degree node for “friends” is “ride” (20 links）and “Disneyland” (20 links). The 20 concepts they related to are mainly describing the intuitive feeling of touring experience, such as “magical”, “worth” and “amazing”. “Park” is the median node, and it is also related to the entertainment facilities. Besides, there also exist important nodes such as “time” and “food”. Thus, to the group of “friends”, the focus point for them are the intuitive experience feelings during the process of touring in the theme park. A detailed exploration of this result suggests that this is mainly related to the facilities, relating to outward, attribute driven experiences (Morgan and Xu, 2009).

PFNET-Couple:

At the first glance, the highest degree node and median nodes for “couples” are almost the same for the group of “friends”. However, the related nodes connected to the highest degree node for “couples” mainly describe the personal psychological feelings, like “special”, “attraction” and “feel”. Similarly, the nodes connected to the median nodes here are focused on the psychological feelings, such as “happy”, “love”, “help” and “memory”. Therefore, to the group of “couples”, the focus point are concentrated on the inner feelings of the commenters themselves (or their partners). This is related to their personal emotional feelings about the experiences, referring to the inward driven experiences (Morgan and Xu 2009). It is possible, when travelling as a couple, they pay more attention to their emotional feelings and memories of the experiences, rather than other aspects.

PFNET-Family:

Compared to the former two spanning trees, there exists one special high degree node in the tree of “family”, which is the word “time”. Considering that the commenters of this group are mostly the middle-age (compared to the elder and children), the comments they provided contain the problems and points that people would face or concern at all ages, especially for kids and elderly people. In this tree, the words related to the high degree nodes are mostly describing the tour schedule, such as “night”, “space”, “August”, “walked”. Besides, words related to the service attitude, such as “rude” and “nice” are also the relatively centralized nodes. This clearly indicates that this particular group pay more attention to the interaction with staff, the service etc., suggesting an interaction driven experience.

The above results clearly suggest different segments pay attention to different aspects of experiences. And path scaling, as an analytical tool, is able to identify the differences of the groups.

#### Similarity Comparison

Furthermore, a comparison between the 4 PFNETs produced have been conducted in pair. We compared the PFNETs in pair and identified the difference discrepancies accordingly with the use of similarity metrics measurement approach. Network similarity varies from 0 to 1, indicating the proportion of the unique links in two network that are found in both networks. The similarity metric is shown in Eq. (1).

C = number of links common to the networks

L1= number of links in first network

L2 = number of links in second network

Table 2 shows estimates of the similarity between these networks, and the data from which one of these estimates was derived. Each network contained 398 links while the links shared in pair (i.e. common links) are different. PNET-Family and PNET-couple has 68 links shared whereby the links in common between PNET-Friend and Total is only 18.

If the similarity value is 1, it depicts an identical network, and 0 denotes no shared links. For instance, the probability of 22 shared links (between Couple and Total) is 0.028424, indicating that the similarity is not statistically significant.

The similarity of the spanning trees between “family” and “couple”, “family” and “friends”, “friends” and “couple” approximately are 0.05, 0.04 and 0.04. While their high degree nodes, which are presenting the words people mentioned in their comments most times, are roughly the same, the center nodes of them show their specificity. For the spanning tree of “friends”, the center nodes are almost the same as its high degree nodes. This means that people visiting with **friends** value the hardware entertaining facility but have no particular focusing entertaining equipment on their tour, as young people have broad interest points. The center nodes of the tree of “family” is almost the same as the center nodes of the “total” tree. On one hand, people visiting the Disneyland in the group of “family” make up a large proportion of the total visiting population. The data analysis also shows that the similarity of links between “family” and “total” is 0.09, mush higher than the similarity between “friends” and “total” (0.02), and between “couple” and “total” (0.03). Therefore, the experiences of “family” is relatively representative. On the other hand, the center nodes of “family”, such as “mountain” and “space”, shows that the biggest concerns of **family** are safety and comfort level, that is the sensory element and humanistic concern. For the group of “couple”, the center nodes, such as “pirate”, “Caribbean”, “haunted”, indicate that compared to the group “friends”, **couples** particularly focus on one or several specific entertaining equipment. In a nutshell, low similarity between PFNETs of Family, Couple and Friend shows that for better marketing the Disneyland would shape different (specific) touring programs for different groups of people, according to their features outlined above.

### Common links between the groups:

It is a little bit hard to give some generalized results through the common links between the groups, as their features and common characters have already been stated above. As a result, (in my opinion) these three graphs can be compared again to find the specific points (or links) of every “common link” graph, which may be useful for the practical situation. For example, the specific link of the graph “Friends and Couple” is the word “time” linking to “wait” and “little”, which are describing the concrete feelings of time during their tour. While in the other two graphs the word “time” links to some general words describing the season or holiday. Therefore, in the practical situation, when the statistical data shows that on a specific day (or time period) most of the tourist are friends and couple, Disneyland could focus more on how to manage the park and some popular entertaining facilities properly that the waiting time would be less or tolerable for the tourists. In addition, another special link in the “Friends and Couple”: “ride-line” can be another proof of this point of view.

Similarly, the link “park-service” in the graph “Family and Couple”, the link “park-parking” and “time-Christmas” in the graph “Family and Friends” also can be processed by the method above.

# **Findings and Conclusions**

Based on 5-year period online customer reviews available on social media site of TripAdvisor, this paper utilized PFNETs to visualize, compare and analyse online reviews out of the full sample of entire visitor tank and its three subsamples – i.e. the large segments of Disneyland visitors (i.e. Family, Friend, Couple) in Disneyland California. By adopting PFNET methods, we measured and compared the concepts relatedness. The results identified the perception or concept dissimilarity out of the three segments when evaluating theme park visitor experiences. The research finds out that different segments pay attention to different aspects of tourist experiences. People travelling with friends usually focus on the outward attributed driven experiences, people travelling with couples like to focus on the inward psychological driven experiences, while people travelling with family members pay particular attention to interaction driven experiences, such as the interaction with staff members, etc. The results prove the facets of tourist experiences suggested by Morgan and Xu (2009) and Kapferer (1998). In addition, our research also identifies different segment groups would pay attention to different aspects of experiences in theme parks, suggesting it is critical to understand and focus on the differences in segments when marketing.

### *Theoretical and practical implications*

Theoretical contributions:

Customer experiences is an essential part of service quality and marketing. However, previous studies on experiences have mainly focused on a homogeneous market which tends to be generic and ignores the difference in segments. Particularly in tourism related studies, there is an urge on focusing on induvial tourist experiences and emphasis on different segments in experiential marketing (Williams, 2006; Xu et al. 2011). Meanwhile, in the tourism studies, there is a trend of shifting from traditional survey based or interview-based studies to big data analytics, which benefits from unconventional dataset, massive amount of information and real time reflection. Our study has made a substantial contribution to the understanding of the visitor experiences in theme parks, particularly identifying the differences and similarities in different segments. The focus of experiences identified by three different segments proves the three different aspects of tourist experiences, proving Morgan and Xu (2009) and Kapferer’s (1997) experiences facets can be applied in theme parks.

Methodological contributions:

Digital texts, especially Online customer reviews have been ubiquitous in today’s extremely competitive tourism and hospitality industry. To gain insight into customers’ preference and their satisfaction embedded in the digital texts, a critical question arises as to how an individual customer’s comprehensive perception to the business service can be captured. Researchers applied various process methods, such as mind-maps (Clariana, 2010; Gonzalvo et al.). However, the shortcoming of this approach is that, in order to get a comprehensive mind-map, time-consuming and high standard training, practice and how to set up objective mind-map scoring are required prior to drawing a satisfying mind-map (Clariana et al., 2006; Fesel et al., 2015).

Presently, the well-recognized statistical techniques for analysing customer experience and satisfaction are cluster analysis approaches, however the weakness of cluster techniques occurs during the cluster set selection and outlier (noise) detection (Hair et al, 2010; Kumar et al., 2014). Other statistical techniques are based on probability theories but was criticized due to its lack of recognition to linguistic uncertainty (Ip et al., 2012). Some new design-science research approaches were proposed with sophisticated methods (Law et al, 2010), but they were regarded as over complicated (Hao, et al, 2015).

In contrast, our paper is the first theme park study in the theme park industry which applies pathfinder network scaling approach to set a fully connected network which indicate the importance of the terms and their relationship in the perception of different customer segments. The advantage is that this approach allows quantitative analysis of structural knowledge for rapid and objective evaluation. Pathfinder’s visual representation and useful function, such as similarities, ensures the evaluation rapid, objective and comprehensive. In addition, the data we used from TripAdvisor information is publicly available which means that this research approach can be applied for the study of other digital texts which are derived from open sources online.

### *Practical implications:*

The fast development of internet and mobile technology has brought strategic changes of Disney and other theme parks. In a data driven era, how to use big data to provider a more personalised customer experiences is the key to success. Cross-segment experience analysis can help theme park managers to: understand the customer preference better and is useful for optimising the resource allocation; identify which activities are least engaging for visitors to each customer segment and reconsider using different marketing tools to present tangible and intangible products in efficiency.

### *Research limitations*

The study has several limitations, which deserve further research. Future research will consider and assessment the segment difference caused by countries, cultures and regions. Other variables to be considered will include preferences between regular visitors and first-time visitors. On obtaining a thorough database, further studies should consider the use of offline questionnaires and interviews to related theme park visitors whose feedback cannot be submitted online.

# **References:**

* All, A.C., Huycke, L.I., & Fisher, M.J. (2003). Instructional tools for nursing education: Concept maps. *Nursing Education Perspectives*, 24, 311-317.
* Amith, Muhammad, et al. (2017). Using Pathﬁnder networks to discover alignment between expert and consumer conceptual knowledge from online vaccine content. *Journal of Biomedical Informatics*, 74, 33–45.
* Arenas-Gaitan, J., Rondan-Cataluña, F.J. & Ramírez-Correa, P. (2013). Social identity, electronic word-of-mouth and referrals in social network services. *Kybernetes*, 42(8), 1149-1165.
* Birenboim, A., Anton-Clavé, S., Russo, A. P., & Shoval, N. (2013). Temporal activity patterns of theme park visitors. *Tourism Geographies*,15 (4),601–619.
* Bitner, M. J. (1992). Servicescapes: The Impact of Physical Surroundings on Customers and Employees. *Journal of Marketing*, 56(2), 57-71.
* Brown, T.J., Churchill, G.A. Jr & Peter, J.P. (1993). Improving the measurement of service quality. *Journal of Retailing*, 69 (1), 127-139.
* Chiappa, G. D., Ladu, MG., Meleddu, M., & Pulina, M. (2013). Investigating the degree of visitors’ satisfaction at a museum. Anatolia. *An International Journal of Tourism and Hospitality Research*, 24(1), 52–62.
* Clariana, R. B., Koul, R., & Salehi, R. (2006). The criterion-related validity of a computer-based approach for scoring concept maps*. International Journal of Instructional Media*, 33(3), 317–325.
* Clariana, R. B. (2010). Deriving group knowledge structure from semantic maps and from essays. In D. Ifenthaler, P. Pirnay-Dummer, & N. M. Seel (Eds.), Computer-based diagnostics and systematic analysis of knowledge (pp. 143–159). New York, NY: Springer. Chapter 7.
* Crick, J., & Crick, D. (2016). Developing entrepreneurial resilience in the UK tourism sector. *Strategic Change: Briefings in Entrepreneurial Finance,* 25, 315-325.
* Crosby, L.A., Evans, K.R. & Cowles, D. (1990). Relationship quality in services selling: an interpersonal inﬂuence perspective. *Journal of Marketing*, 54 (3), 68-81.
* Day, E.A., Arthur, W., & Gettman, D. (2001). Knowledge structures and the acquisition of a complex skill. *Journal of Applied Psychology*, 86, 1022-1033.
* Echtner, C.  Ritchie, J. B. R. (2003). The meaning and measurement of destination image. *Journal of Tourism Studies*, 14(1), 37–48.
* Felbermayr, A., & Nanopoulos, A. (2016). The Role of Emotions for the Perceived Usefulness in Online Customer Reviews. *Journal of Interactive Marketing*, 36, 60-76.
* Fesel, S.S., Segers, E., Clariana, R.B., & Verhoeven, L. (2015). Quality of children’s knowledge representations in digital text comprehension: Evidence from pathfinder networks. *Computers in Human Behavior*, 48, 135-146.
* Fotiadis, A.K. (2016). Modifying and applying time and cost blocks: the case of E-da theme park, Kaohsiung, Taiwan. *Tourism Management*, 54,34-42.
* Fotiadis, A., and Kozak, M (2017). Managing the perception of service quality; the importance of understanding differences between demographic and behavioural customer segments amongst theme park visitors. *Facilities*, 35 (9/10),.486-510.
* Freytag, P., & Clarke, A. (2001). Business to business market segmentation. *Industrial Marketing Management*, 30,473- 486.
* Frochot, I., & Morrison, A. M. (2000). Beneﬁt segmentation: A review of its applications to travel and tourism research. *Journal of Travel & Tourism Marketing,* 9(4), 21-45.
* Geetha, M., Singha, P., & Sinha, S. (2017). Relationship between customer sentiment and online customer ratings for hotels - An empirical analysis. *Tourism Management*, 61, 43-54.
* Geissler, G. L., & Ruck, C.T. (2011). The overall theme park experience: A visitor satisfaction tracking study. *Journal of Vacation Marketing*, 17(2), 127–138.
* Goldsmith, T., Johnson, P., & Acton, W. (1991). Assessing structural knowledge. *Journal of Educational Psychology*, 83(1), 88–96.
* Gonzalvo, P., Cañas, J. J., & Bajo, M. (1994). Structural representations in knowledge acquisition. *Journal of Educational Psychology*, 86(4), 601–616.
* Hair, J., Black, W., Babin, B., & Anderson, R. (2010). *Multivariate data analysis* (7th ed.). England: Pearson Prentice Hall.
* Hao, JX., Yu, Y., Law R., and Fong D. (2015). A genetic algorithm-based learning approach to understand customer satisfaction with OTA websites. *Tourism Management.48,* 231-241.
* Hennig-Thurau, T., & Klee, A. (1997). The impact of customer satisfaction and relationship quality on customer retention: a critical reassessment and model development. *Psychology and Marketing*, 14(8), 737-764.
* Ho-Dac, N. N., Carson, S. J., & Moore, W. L. (2013). The effects of positive and negative online customer reviews: Do brand strength and category maturity matter?. *Journal of Marketing*, 77(6), 37-53.
* Homburg, C., Koschate, N., & Hoyer, D. (2005) Do Satisfied Customers Really Pay More? A Study of the Relationship Between Customer Satisfaction and Willingness to Pay. *Journal of Marketing*: 69 (2), 84-96.
* Housner, L. D., Gomez, R., & Griffey, D. C. (1993). A pathfinder analysis of pedagogical knowledge structures: A follow-up investigation. *Research Quarterly for Exercise and Sport*, 64(3), 291–299.
* Ip, C., Law, R., & Lee, H. A. (2012). The evaluation of hotel website functionality by fuzzy analytic hierarchy process. *Journal of Travel & Tourism Marketing*, 29(3), 263 -278.
* Jonassen, D. H. (1993). Effects of semantically structured hypertext knowledge bases on users’ knowledge structures. In C. McKnight, A. Dillon, & J. Richardson (Eds*.), Hypertext: A psychological perspective*, 153–162. London: Ellis Horwood.
* Kapferer, J-N. (1997). *Strategic brand management: New approaches to creating and evaluating brand equity*. London: Kogan Page.
* King, J. (2002). Destination marketing organisations – Connecting the experience rather than promoting the place. *Journal of Vacation Marketing*, 8(2), 105-108.
* Kumar, V., Tan, P., & Steinbach, M. (2014). *Introduction to data mining.* USA: Pearson New International Edition.
* Lau, W. and Yuen, A. (2009). Predictive validity of measures of the pathﬁnder scaling algorithm on programming performance: alternative assessment strategy for programming education. *Journal of Educational Computing Research,* 41(2), 227-250.
* Law, R., Qi, S., & Buhalis, D. (2010). Progress in tourism management: a review of website evaluation in tourism research. *Tourism Management*, 31(3), 297-313.
* Martilla, J.A. and James, J.C. (1977). Importance-performance analysis*. Journal of Marketing*, 41(1), 77-79.
* Min, H., Lim, Y., and Magnini, V. P. (2015). Factors Affecting Customer Satisfaction in Responses to Negative Online Hotel Reviews: The Impact of Empathy, Paraphrasing, and Speed. *Cornell Hospitality Quarterly*, 56(2), 223 –231.
* Morgan, M., & Xu, F. (2009). Student travel experiences: Memories and dreams. *Journal of Hospitality Marketing & Management*, 18(2), 216-236.
* Nizamuddin, SK. MD. (2015). Marketing utility of Tripadvisor for hotels: an importance-performance analysis. *Journal of Tourism*, 16 (1), 69–75.
* Oh, H. (2001). Revisiting importance–performance analysis. *Tourism Management*, 22 (6), 617-627.
* Parasuraman, A., Zeithaml, V., & Berry, L. L. (1988). SERVQUAL: A multiple-item scale for measuring con-sumer perceptions of service quality. *Journal of Retailing*, 64, 12–40.
* Pine, B. Joseph & James. H, Gilmore (1999), The experience economy. Boston, MA: Harvard Business School Press.
* Pikkemaat, B., & Schuckert, M. (2007). Success factors of theme parks – an exploratory study. *Tourism*, 55(2), 197-208.
* Prayag, G., & Hosany, S. (2014). When Middle East Meets West: Understanding the Motives and Perceptions of Young Tourists from United Arab Emirates. *Tourism Management*, 40, 35-45.
* Prayag, G., Disegna, M., Cohen, S. A., & Yan, H. G. (2015). Segmenting markets by bagged clustering young chinese travelers to Western Europe. *Journal of Travel Research*, 54(2), 234-250.
* Rau, P.-L., Gao, Q. and Ding, Y. (2008). Relationship between the level of intimacy and lurking in online social network services. *Computers in Human Behavior*, 24(6), 2757-2770.
* Rashidirad, M., Salimian., H., Soltani, E., & Fazeli., Z. (2017). Competitive strategy, dynamic capability, and value creation: Some empirical evidence from UK telecommunications firms. *Strategic Change: Briefings in Entrepreneurial Finance,* 26(4), 333-342.
* Roblek, V., Pejic Bach, M., Meško, M.& Bertoncelj, A. (2013). The impact of social media to value added in knowledge-based industries. *Kybernetes*, 42(4), 554-568.
* Sánchez-Franco M.J., Navarro-García A., & Rondán-Cataluña F.J. (2016). Online customer service reviews in urban hotels: A data mining approach. *Psychology & Marketing*, 33(12), 1174-1186.
* Sánchez-Franco, Manuel J., Muñoz-Expósito, M., and Villarejo-Ramos, A F. 2017. A knowledge structures exploration on social network sites. *Kybernetes*, 46(5) ,818-839.
* Sievänen, T., Neuvonen, M., & Pouta, E., (2011). National park visitor segments and their interest in rural tourism services and intention to revisit. *Scandinavian Journal of Hospitality and Tourism*, 11, 54-73.
* Smith, W. (1956). Product differentiation and market segmentation as alternative marketing strategies. *Journal of Marketing*, 21(July), 3-8.
* Slak Valek, N., Shaw, M. & Bednarik, J. (2014). Socio-demographic characteristics affecting sport tourism choices: a structural model. *Acta Gymnica*, 44 (1), 57-65.
* Tasci, A. D.A. & Milman A. (2019) Exploring experiential consumption dimensions in the theme park context. *Current Issues in Tourism*, 22(7), 853-876.
* TEA/AECOM. (2016) *TEA/AECOM 2016* *Theme Index and Museum Index: The Global Attractions Attendance Report*. Publisher: Themed Entertainment Association (TEA). Available at: <http://www.teaconnect.org/images/files/TEA_235_103719_170601.pdf> (Accessed: 30, Dec 2018)
* TripAdvisor Source of data link: <https://www.tripadvisor.co.uk/Attraction_Review-g29092-d103346-Reviews-or320-Disneyland_Park-Anaheim_California.html>
* Tsang, N.K.F., Lee, L.Y.S., Wong, A. & Chong, R. (2012). THEMEQUAL – adapting the SERVQUAL scale to theme park services: a case of Hong Kong Disneyland. *Journal of Travel & Tourism Marketing*, 29(6), 416-429.
* Van Ree, H.J. (2009). *Service Quality Indicators for Business Support Services*. University College, London.
* Vassiliadis, C.A., Priporas, C.-V. and Andronikidis, A. (2013). An analysis of visitor behaviour using time blocks: a study of ski destinations in Greece. *Tourism Management,* 24, 61-70.
* Wang, J.C., & Chang, C.-H. (2013). How online social ties and product-related risks influence purchase intentions: a facebook experiment. *Electronic Commerce Research and Applications,* 12 (5), 337-346.
* Williams, M. (2004). Concept mapping: A strategy for assessment. *Nursing Standard*, 19(9), 33-38.
* Williams, A. (2006). Tourism and hospitality marketing: Fantasy, feeling and fun. *International Journal of Contemporary Hospitality Management*, 18, 482–495.
* Xu, F; Morgan, M; Moital, M. (2011) Cross Cultural Segments in International Students Travel—An analysis of the British and Chinese Market. *Tourism Analysis* 16(6), 663-675.
* Xiang, Z., Schwartz, Z., Gerdes, J. H., Jr, & Uysal, M. (2015). What can big data and text analytics tell us about hotel guest experience and satisfaction? *International Journal of Hospitality Management*, 44, 120–130.
* Xu, F., Morgan, M., Moital, M. (2011). [Cross-cultural segments in international student travel: an analysis of British and Chinese market](https://www.ingentaconnect.com/content/cog/ta/2011/00000016/00000006/art00003). *Tourism Analysis*, 16(6), 663-675.
* Zeithaml, V.A., Berry, L.L. & Parasuraman, A. (1988). Servqual: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing*, 64, 12-40.

# Tables and Figures

Table 1: TF-IDF high frequency word list (shown in part)

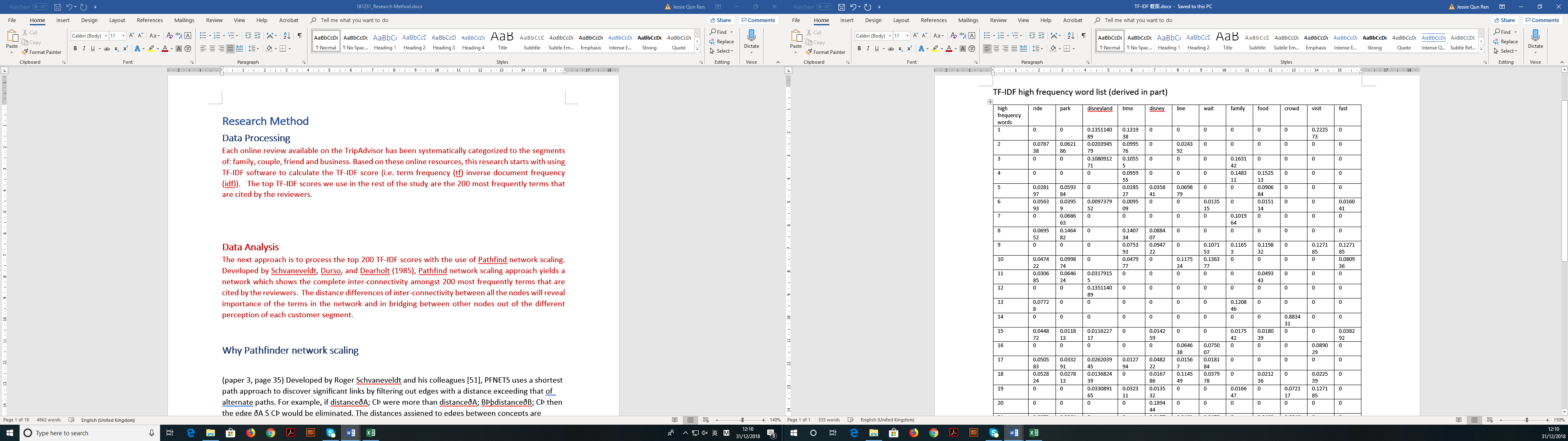


Table 2: Combined PFNETs similarity Comparison

|  |  |  |  |
| --- | --- | --- | --- |
| Similarity of PNET  links between each pair | Number of PNET links shared in pair | Number of Links in each individual PFNET (See figure a, b, c, d) | Similarity Value  (PNETs in pair) |
| **Family and Couple** | 38 | 398 | 0.050132 |
| **Family and Friend** | 30 | 398 | 0.039164 |
| **Friends and Couple** | 34 | 398 | 0.044619 |
| **Friends and Total** | 18 | 398 | 0.023136 |
| **Family and Total** | 68 | 398 | 0.093407 |
| **Couple and Total** | 22 | 398 | 0.028424 |

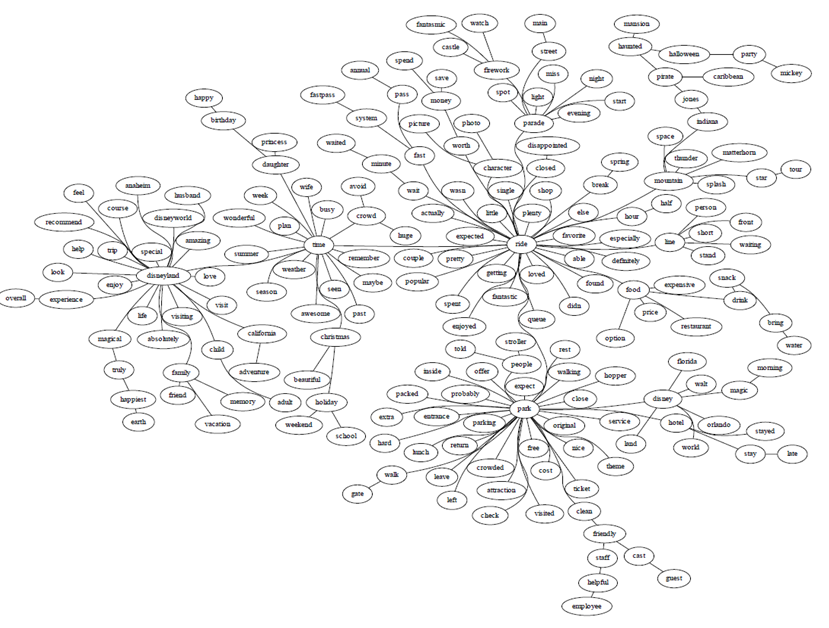


Figure 1: PFNET- Total

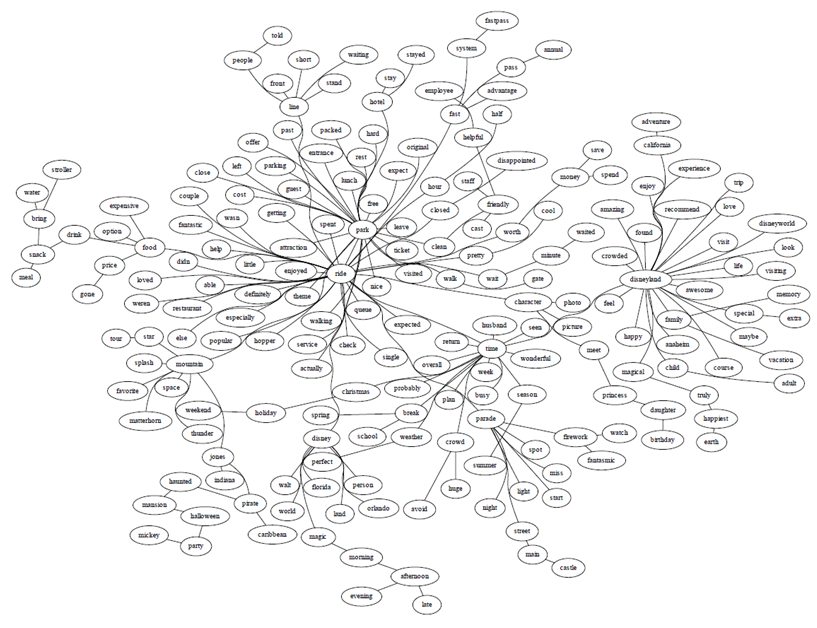


Figure 2: PFNET- Family

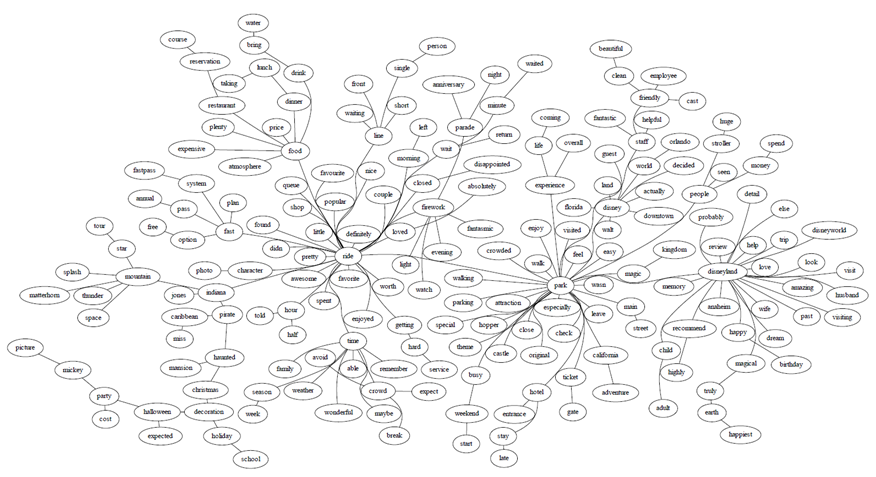


Figure 3: PFNET- Couple

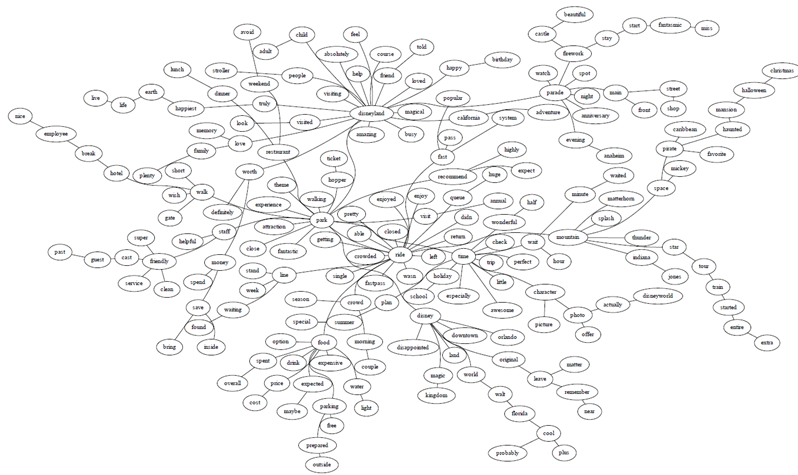


Figure 4: PFNET- Friend