Probing the Interconnections Between Geo-Exploration and Information Exploration Behavior

Dongho Choi, Chirag Shah, Vivek Singh
School of Communication & Information (SC&I)
Rutgers, The State University of New Jersey
4 Huntinton St, New Brunswick, NJ 08901
{dongho.j.choi,chirags,vivek.k.singh}@rutgers.edu

ABSTRACT

As increasingly diverse facets of human life - including socializing, exercising, and information-seeking - are mediated by ubiquitous technology, they open the doors for the study of the hitherto under-explored interconnections between them. This work motivates and grounds the use of geo-exploration data to predict the information exploration behavior of users and to support their search. Based on a two-week field study involving 35 participants, we have identified multiple geo-exploration features that have significant associations with a user's information exploration behavior. We also found that the same geo-exploration features could be combined to build predictive models for various facets of an individual's information exploration behavior, and these models performed significantly better than comparable personality-based models.

Keywords

Geo-exploration; Information exploration; Exploratory search

Categories and Subject Descriptors

H.3.3 [**INFORMATION STORAGE AND RETRIEVAL**]: Information Search and Retrieval—*Search process*

1. INTRODUCTION

This work studies the interconnections between individuals' geographic exploration and their information exploration behavior. Such an interconnection between geographic and online information exploration has been alluded to in the past via conceptual metaphors like information "foraging" and through recent empirical studies like the one by Barbosa et al.[2], which has found a dichotomy of *explorers* and *returners* to hold in both online and geographic settings [31, 29, 2]. While Barbosa et al. [2] focus on broad consistencies in population level metrics, they did not study whether the **same** individuals tend to be explorers in both geographic and online settings or if a combination of geo-exploration features can predict an individual's information exploration behavior.

UbiComp '16, September 12 - 16, 2016, Heidelberg, Germany

ISBN 978-1-4503-4461-6/16/09...$15.00

DOI: http://dx.doi.org/10.1145/2971648.2971694

The propensity of an individual to obtain the right information from a search engine is based not just on the search results, but also the way the user interacts with the results provided. Some users tend to be "fast surfers" while others tend to be "deep divers" [19]. Consequently their information outcomes can be different, affecting multiple facets of their lives including health management, financial decisions, and job performance [6, 27]. Multiple researchers have proposed different methods to nudge users of different personality types and predispositions towards better information outcomes [45, 26]. For example, pre-emptively identifying that a user is likely to under-explore the search results could be used to tailor the search results accordingly [14], change the result interface, or design interventions [36] that nudge the user into exploring more web pages before finishing their search session.

The majority of work on pre-identifying user behavior has focused either on explicit questionnaires, past search information, or personal characteristics such as socio-demographic descriptors and personality types [10]. In this work, we push this direction further and investigate if an individual's geo-exploration signature can predict aspects of information behavior, and perhaps, perform at par or even better than personality descriptors at doing so. Such an interconnection between geo-exploration and information exploration, if found, will open a two-way door allowing the use of a whole new category of data coming from geo/mobile sensors for better information exploration, and also for using online information exploration behavior for recommending geographic travel paths and destinations.

2. BACKGROUND

White and Roth [45] propose a set of features that must be presented in exploratory search systems, which includes (1) support querying and rapid query refinement, and (2) leverage search context. Because user-generated queries may be based on their existing knowledge and impose limitations on exploratory search, the presentation of query suggestions helps users explore additional query terms. For example, dynamic queries method could be employed for such rapid query refinement [1]. Dynamic queries method provides users with ways of seeing an overview of the data, exploring it rapidly, and filtering out unwanted information. Through adjusting a query (with sliders, switches, and filters) users explore the information space, and view the changing results. In the meantime, contextual information of users can be used to automatically suggesting the right parameters for the various sliders and switches. Besides the context captured by explicitly asking users or implicitly by mining users' past interaction behavior [11, 21], geo-exploration data as described in this work could also be used to define such a personalized model of search behavior.

When it comes to the behavioral models, past studies have acknowledged the effect of individual personalities, demographic descriptors, and personal contexts in human information behavior (e.g., [4]), but most studies have focused on traits that could be easily observed (e.g. gender, ethnicity) or elicited in a short time in laboratory settings. Unfortunately, the personal contexts captured by observations in limited, unnatural settings must contend with multiple challenges, including subjectivity in observation, biases, and limited observation opportunities [12, 15, 32]. On the other hand, mobile phones and sensors in people's daily living help researchers build personalized models of human behavior in different contexts and connect them to other outcomes such as health, financial outcomes, depression, happiness, and academic accomplishments [38, 23, 5, 47, 41, 42, 44].

An interconnection between geo-exploration and information exploration has been alluded to in multiple recent studies. While Hills et al. [20] review the exploration behaviors revealed in different contexts (e.g. foraging, visual search, social search, etc.), a recent study with fMRI experiments found that people use similar brain structures during both navigating behaviors over physical space and digital space [3].

Specifically, while the previous research has reported broad population level similarities between human mobility and web exploration behavior, we investigate this at an individual level, and study the following research question:

Can a specific aspect of human information behavior - exploration in search - be predicted by the same person's geo-exploration behavior?

3. METHODOLOGY

To investigate how people's geo-exploration behaviors relate to their search exploration behaviors, we conducted a hybrid study with a field study and a lab experiment. This section provides details about the design of the study.

3.1 Participants

In the recruitment stage for this study, we asked students of Rutgers University to participate in a study through open calls via various email-lists and social networking sites (e.g., Facebook). Participants were each compensated $40 in cash upon completion of the study. A total of 35 students participated during spring and summer semesters of 2015.

3.2 Session Workflow

Each session's workflow for this study is described in Table 1. We invited participants to a lab in the school to introduce the study, and install required apps on their phone. During the field study, participants were not asked to do anything specific except keep the mobile app installed and answer the Big Five personality questionnaire [16].

3.3 Mobile App

The participants were given an app for their smartphones (Android OS) that recorded data about different locations they visit as well as their phone call and SMS usage data. The app captured GPS based location as well as approximate location based on nearest cell tower. The unit of analysis for location data in this study was the latitude, longitude tuple at the fourth decimal point resolution. This roughly corresponds to 10m by 10m blocks and was selected as the scale since current GPS data have a pseudo-range accuracy of about 8 meters at 95% confidence level[1]. All data collected through the mobile app were anonymized.

[1] http://www.gps.gov/systems/gps/performance/accuracy/

3.4 Lab Session

At the end of two-week field session, we brought the participants to the lab where they were given an exploratory search task on the topic of 'cyber-security' for 30 minutes, collecting as much information as possible to write a report as a result of the search. During the task, users' actions in the browser, such as visited Webpages and queries, were recorded with timestamps. The task was presented as follows:

"You are a journalist for a national newspaper. With the recent focus on security breaches and consumer data breaches, vulnerabilities in data and software has become an important topic of interest to a multitude of people ranging from government, companies, technical industry to general public.

You have been asked to write an article about data and software vulnerabilities. Your article should be around 1000 words. Your article should cover basic introductions to the topic, along with past and present events. Your article should appeal to as many people as possible, including an unfamiliar, lay audience and people in affected businesses. It should cover different aspects of software vulnerabilities that are relevant to their daily lives. The article should also focus on the measures taken to minimize these risks at industry, governments and consumer levels.

When conducting this task, you should collect as much information as you can by searching online. You can use snippets to collect information that you deem useful for writing the report and copy them into your article. Nevertheless, the article should be well-written with a proper flow so that the readers would be able to understand the context and get lots of information about this topic."

4. EVALUATION FRAMEWORK

This section presents our evaluation framework that characterizes exploratory behavior in search and in physical movment.

4.1 Exploration Behavior in Search

Most information seeking models [9, 13, 22] characterize exploration behavior in search tasks using the act of searching (through queries, questions), and the outcome of searching (sources, documents). This insight provides us with two aspects of search tasks where we could investigate exploration behavior: *input to search as expressed by queries*, and *output of search as represented by the nature of the information (Webpages, in this case) discovered i.e. clicked*. The former can be measured using inter-query and within-query diversity and the latter can be measured using novelty of discovery suggested in [33]. These three measures are explained below in details.

4.1.1 Inter-Query Diversity (ID)

One indicator of exploratory behavior in the searches conducted by a user is the difference in attempted queries. To find the difference between two queries, we used *Generalized Levenshtein (edit) distance*, a commonly used distance metric for measuring the distance between two character sequences [24].

4.1.2 Within-Query Diversity (WD)

In information theory, *entropy* [34] refers to the amount of uncertainty of an unknown or random quantity. In information retrieval literature, entropy is used to measure the effectiveness of combination of terms, for example, via the Maximum Entropy Principle (MEP) [8]. Hence, here we gauge the relative unique and informative value of a specific query with an entropy-based measure. We

Session	Procedure	Description	Time
Field Session	Introduction	We introduce the study, and install the required app.	30 mins
	Field task	Participants continue to use their phones throughout their everyday lives as the app collects their personal and contextual signals. In the middle of this session, they are asked to answer to the personality questionnaire.	2 weeks
Lab Session	Introduction	We introduce the lab session and information-seeking tasks the participants will be given.	10 mins
	Lab Task	Participants individually conduct an exploratory search task, including pre-survey and post-survey.	30 mins
	Wrap-up	Wrap-up and (optional) interview.	30 mins

Table 1: Session Workflow

first calculate the information entropy as follows:

$$Entropy_{Q_a} = -\sum_{i=1}^{|unigrams_{Q_a}|} p_u log_2 p_u \quad (1)$$

Where p_u is the frequency of counts of each unigram, u appearing in each query string, Q_a, found in the entire dataset. The within-query diversity for each user can be defined as the mean of entropy values of each distinct issued query as follows.

$$WD = \frac{\sum_{a=1}^{NQ} Entropy_{Q_a}}{|Q|} \quad (2)$$

Higher value of within-query diversity means that the participant has entered diverse keywords in a query during the search task.

4.1.3 Novelty of Discovery (ND)

In an exploratory search task where there is plenty of information available for a topic, one does not need to be skilled at exploration to find a large amount of information. Therefore, what really exhibits one's exploratory behavior is the ability to find information that is novel - not found by many. To measure this, we used Likelihood of Discovery (LD), defined by Shah and González-Ibáñez [33]. LD for webpage wp_i is defined as follows:

$$LD_{wp_i} = \frac{-1 \cdot n\{wp_i\}}{|U|} \quad (3)$$

Here, $n\{wp_i\}$ is the number of users who found wp_i and $|U|$ is the total number of users. Therefore, LD for a page goes from $-1/|U|$ (highest) to -1 (lowest). The Novelty of Discovery measure for each user can be found as follows:

$$ND = \frac{\sum_{i=1}^{|Coverage|} LD_{wp_i}}{|Coverage|} \quad (4)$$

Where $|Coverage|$ is the set of distinct content pages that a user visited. Thus, ND goes from $-1/|U|$ (found most novel pages) to -1 (found most common pages). For a large N this metric ranges between -1 and 0 with a higher value indicating a higher propensity to find novel information.

4.2 Modeling Geographic Exploration

Animal foraging literature has shown that different animals exhibit individualistic patterns of foraging (e.g. [18]). When presented with a spatial distribution of "patches" with different utilities, animals face a trade-off between the conflicting demands of sampling a variable environment and the exploitation of the most profitable resources (e.g. [7]). The response of different animals to this trade-off varies, and while some animals demonstrate higher degrees of "diversity" across patches, others tend to show higher degrees of "loyalty" to the known patches [39, 35, 17, 46]. We posit that similar traits in human mobility behavior may be useful to model human mobility in the information exploration context. Note that here we considered the location data from the whole field session period.

4.2.1 Unique Locations

This is a measure of location-based activities: the total number of unique locations visited by a user during the field session.

4.2.2 Location Diversity

$$D = -\sum_i p_i \cdot log_b(p_i) \quad (5)$$

Where p_i = percentage of overall visits that were devoted to location i, and b is the total number of unique locations visited.

The diversity score measures how evenly a user's geographic movements are distributed between different locations, using Shannon Entropy [40, 28]. A user with low diversity distributes her time unevenly across locations, whereas a user with high diversity spends time evenly across many locations.

4.2.3 Location Loyalty

Loyalty characterizes the repetitions in user movements. Similar to an animal going back to the same patch of grass, this characterizes the tendency of a user to go back to his favorite locations.

$$L = \sum_i^k p_i \quad (6)$$

Where p_i = percentage of overall visits that were devoted to location i, and k is the chosen threshold on a user's favorite locations. In this work, we consider k to be top-third of a given user's most frequent locations. Note that we employed the meaning of loyalty used in [37] and adjusted it to our context given the limited number of locations visited by the participants. A similar top-third threshold has been adopted by multiple studies to demarcate the groups or the relationships with highest tie strength or social prestige in the past (e.g. [25].) In essence, the loyalty feature captures the degree of repetition in a participant's movement patterns.

The resulting values are between 0 and 1, with the larger numbers meaning higher spatial loyalty. For example, a user with very high spatial loyalty will spend almost all of her time (e.g. 80%) in her favorite locations, while a user with low spatial loyalty might spend only 40% of her time in the top-third of her visited locations.

5. RESULTS

This section presents our findings resulting from exploratory analysis. This analysis starts with examining relationships between variables using correlation analysis, and proceeds to build predictive models. Please note that for analysis we used data from 31 participants since geo-locational data from four participants were missing.

5.1 Correlation Analysis

Beginning our investigations, we first used the evaluation framework described earlier and looked at if and how different aspects of geo-exploration relate to aspects of search behaviors that exhibit exploration. The results are presented in Table 2.

5.1.1 Unique Location

While Pappalardo et al.'s study [28] found aggregated mobility volume of individuals in municipal level related to the socio-economic developments, a similarly defined feature was not found to have predictive power on information seeking behavior at an individual level.

5.1.2 Location Diversity

Results here show that Location Diversity was positively correlated with Within-Query Diversity (WD), while negatively related to Novelty of Discovery (ND). This indicates that those who visited more diverse locations were also likely to have more within-query diversity and find less unique content. (See Discussion section for more comments.)

5.1.3 Location Loyalty

Location Loyalty was positively correlated with Inter-Query Diversity (ID) and ND, but negatively correlated with WD. This shows that people with higher loyalty to locations were likely to try different queries and encounter more novel information, but they were not quite "adventurous" with each query they issued. In other words, the more loyal someone to a location, the less likely they are to use uncommon terms in their queries.

Variables	Inter-Query Diversity (ID)	Within-Query Diversity (WD)	Novelty of Discovery (ND)
Unique Locations	.285	-.318	.263
Location Diversity	-.261	.433*	-.460**
Location Loyalty	.401*	-.440*	.548**

Table 2: Correlation between Variables (*: $p < 0.05$, **: $p < 0.01$).

5.2 Prediction Model

Multiple significant associations found in the correlation analysis motivated the use of a combination of geo-exploration features to build a predictive model for different facets of information exploration - Novelty of Discovery, Inter-Query Diversity, and Within-Query Diversity - as described in the previous section.

We proceeded to create two classes for the three output variables (*Novelty of Discovery, Inter-Query Diversity, and Within-Query Diversity*) based on a split at the median values. Thus for each variable, we ended up with two equitably sized classes corresponding to "High" (above median) levels of exploration, and "Low" (below median) levels of exploration.

We employed the Naive Bayes classifier to build the classification models and employed 10-fold cross-validation to obtain the results.

We compare the proposed geo-exploration feature-based approach with two different approaches. One is a baseline 'Zero-R' approach, which simply classifies all data into the largest category. The second approach is based on using personality variables, which have been shown by multiple efforts to be related to information exploration (e.g. [43, 19]).

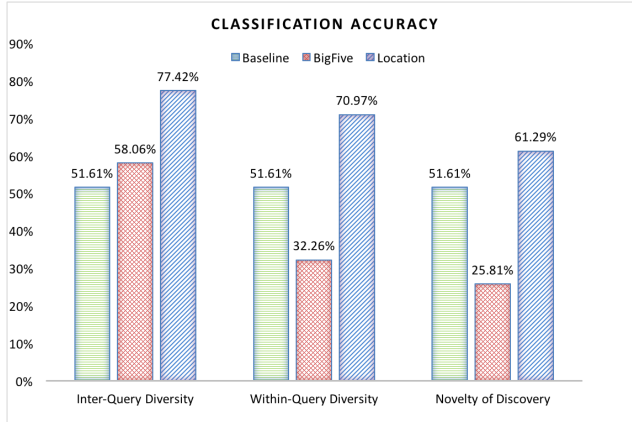

Figure 1: Correct classification percentage for the 3 information exploration behavior features.

Given the equitably distributed classes, the baseline approach simply resulted in 52% accuracy for all the classification tasks. As can be seen in Figure 1, the geo-exploration features perform much better than baseline methods for classification of the information exploration behavior. On average the geo-exploration features could correctly classify the information exploration category 70% of the time, which marks a 35% relative improvement over the baseline method.

The second alternative approach considered was based on personality-based features. Using the well-accepted Big-Five features [16] to represent the personality, the model classified the three output variables correctly 39% of the time. This suggests that traditional personality type measures may not be suited to predict information exploration in short-term exploratory search tasks. At the same time, these results suggest that geo-exploration feature-based models might be able to capture behavioral traits that go beyond the personality variables when explaining the information exploration. Further investigation with larger samples is needed to confirm this initial evidence.

6. DISCUSSION

As suggested in [2], the results of this study also find a dichotomic characteristic between location diversity and location loyalty. In terms of our study's approach, returners can be understood to have higher location loyalty while explorers correspond to higher location diversity. Location diversity and location loyalty in Table 2 consistently encounter opposite directions of associations with other variables, thus suggesting that geograpahic explorers and returners are likely to obtain opposite outcomes in terms of information seeking too.

We also found a dichotomy between two aspects of an user's information behavior: *input to search as expressed by queries* and *output of search represented in the discovered (i.e. clicked) information*. Seeing location diversity and location loyalty, diversity in search keywords (WD) is negatively related to the propensity to find novel information during search (ND). We consider this observation quite interesting and will investigate it further in our future work.

Along the same line of (geo) explorers [29] and Web Explorers [2], diversity in geographical visits (location diversity) is revealed to have positive relationship with diversity of keywords during a search session (WD). This result indicates that a person who explores more places than others might have varied perspectives toward the given topics and/or want to look at diverse aspects of the question.

While previous studies (e.g [43, 19]) have suggested associations between personality traits and information seeking behavior, our results suggest that geo-exploration behavior is able to predict the information-exploration patterns more accurately than personality descriptors. While our methodology is significantly different from previous studies and focused on specific aspects of information-seeking behavior (exploration in short-term search settings), we consider this to be an interesting opportunity for further research.

With further validation, the proposed approach can be used to leverage search context effectively via data from ubiquitous sensors such as smartphones. In addition to search engines adapting the results based on users' online behavioral data, and contextualizing it with immediate sensor metadata (e.g. geo-location), IR systems will be able to support personalized query recommendation methods [45]. For example, dynamic query interface (e.g. [1]), in which slide bars adjust to user's interests and actions, can automatically represent the user's preference to diverse search queries based on the personalized behavioral model. In this sense, while Pirolli and Card [30] attempted to design information systems that provide "information scent" to users so that they could find useful clues during tasks, our findings can guide the system to customize individual "scent" to different users.

7. CONCLUSION

This work motivates and grounds the use of geo-exploration data as obtained via smartphones to understand and predict the information exploration behavior of users. We have identified multiple geo-exploration features that have significant associations with a user's information exploration behavior, and found that the same geo-exploration features can be combined to build predictive models of an individual's information exploration behavior that perform significantly better than personality-based models.

Limitations of our work include the small, relatively homogeneous sample, as well as the constrained setting of undertaking information exploration in a short-term laboratory experiment. Despite the aforementioned limitations, we consider the work to be potentially transformative, as it connects the field of personalized information exploration with human geo-exploration literature. Besides epistemological insights, such a connection may in future lead to use billions of smartphones to design better-suited information experiences for users.

8. ACKNOWLEDGEMENTS

The authors would like to thank Serife Uzun for helping to conduct the study.

9. REFERENCES

[1] Christopher Ahlberg, Christopher Williamson, and Ben Shneiderman. 1992. Dynamic queries for information exploration: An implementation and evaluation. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 619–626.

[2] Hugo S Barbosa, Fernando B de Lima Neto, Alexandre Evsukoff, and Ronaldo Menezes. 2016. Returners and Explorers Dichotomy in Web Browsing Behavior - A Human Mobility Approach. In *Complex Networks VII*. Springer, 173–184.

[3] Yael Benn, Ofer Bergman, Liv Glazer, Paris Arent, Iain D Wilkinson, Rosemary Varley, and Steve Whittaker. 2015. Navigating through digital folders uses the same brain structures as real world navigation. *Scientific reports* 5 (2015).

[4] Tracey A Burdick. 1996. Success and Diversity in Information Seeking: Gender and the Information Search Styles Model. *School Library Media Quarterly* 25, 1 (1996), 19–26.

[5] Michelle Nicole Burns, Mark Begale, Jennifer Duffecy, Darren Gergle, Chris J Karr, Emily Giangrande, and David C Mohr. 2011. Harnessing context sensing to develop a mobile intervention for depression. *Journal of medical Internet research* 13, 3 (Jan. 2011), e55.

[6] Donald Owen Case. 2012. *Looking for information: A survey of research on information seeking, needs and behavior*. Emerald Group Publishing.

[7] Anne Barrett Clark and Timothy J Ehlinger. 1987. Pattern and adaptation in individual behavioral differences. In *Perspectives in ethology*. Springer, 1–47.

[8] William S Cooper. 1983. Exploiting the maximum entropy principle to increase retrieval effectiveness. *Journal of the American Society for Information Science* 34, 1 (1983), 31–39.

[9] Brenda Dervin. 1977. Useful Theory for Librarianship: Communication, Not Information. *Drexel library quarterly* 13, 3 (1977), 16–32.

[10] George A Donohue, Phillip J Tichenor, and Clarice N Olien. 1975. Mass Media and the Knowledge Gap A Hypothesis Reconsidered. *Communication Research* 2, 1 (1975), 3–23.

[11] Susan Dumais, Edward Cutrell, Raman Sarin, and Eric Horvitz. 2004. Implicit queries (IQ) for contextualized search. In *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, 594–594.

[12] Nathan Eagle and Alex Pentland. 2006. Reality mining: sensing complex social systems. *Personal and ubiquitous computing* 10, 4 (2006), 255–268.

[13] David Ellis, Deborah Cox, and Katherine Hall. 1993. A comparison of the information seeking patterns of researchers in the physical and social sciences. *Journal of documentation* 49, 4 (1993), 356–369.

[14] Maria Gäde, Mark Hall, Hugo Huurdeman, Jaap Kamps, Marijn Koolen, Mette Skov, Elaine Toms, and David Walsh. 2015. Supporting Complex Search Tasks. In *Advances in Information Retrieval*. Springer, 841–844.

[15] Jim Giles. 2012. Making the links. *Nature* 488, 7412 (2012), 448–450.

[16] Samuel D Gosling, Peter J Rentfrow, and William B Swann. 2003. A very brief measure of the Big-Five personality domains. *Journal of Research in personality* 37, 6 (2003), 504–528.

[17] Ton GG Groothuis and Claudio Carere. 2005. Avian personalities: characterization and epigenesis. *Neuroscience & Biobehavioral Reviews* 29, 1 (2005), 137–150.

[18] Kristen Hawkes and James F O'Connell. 1985. Optimal foraging models and the case of the! Kung. *American Anthropologist* 87, 2 (1985), 401–405.

[19] Jannica Heinström. 2005. Fast surfing, broad scanning and deep diving: The influence of personality and study approach on students' information-seeking behavior. *Journal of documentation* 61, 2 (2005), 228–247.

[20] Thomas T Hills, Peter M Todd, David Lazer, A David Redish, Iain D Couzin, Cognitive Search Research Group, and others. 2015. Exploration versus exploitation in space, mind, and

society. *Trends in cognitive sciences* 19, 1 (2015), 46–54.

[21] Diane Kelly and Nicholas J Belkin. 2004. Display time as implicit feedback: understanding task effects. In *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, 377–384.

[22] Carol Collier Kuhlthau. 2004. *Seeking meaning: A process approach to library and information services*. Libraries Unltd Incorporated.

[23] Nicholas D. Lane, Emiliano Miluzzo, Hong Lu, Daniel Peebles, Tanzeem Choudhury, and Andrew T. Campbell. 2010. A survey of mobile phone sensing. *IEEE Communications Magazine* 48, 9 (2010), 140–150.

[24] Vladimir I Levenshtein. 1966. Binary codes capable of correcting deletions, insertions, and reversals. In *Soviet physics doklady*, Vol. 10. 707–710.

[25] Ronald Lippitt and Martin Gold. 1959. Classroom social structure as a mental health problem. *Journal of Social Issues* 15, 1 (1959), 40–49.

[26] Gary Marchionini. 1989. Information-seeking strategies of novices using a full-text electronic encyclopedia. *Journal of the American Society for Information Science* 40, 1 (1989), 54–66.

[27] Elizabeth Wolfe Morrison. 1993. Newcomer information seeking: Exploring types, modes, sources, and outcomes. *Academy of management Journal* 36, 3 (1993), 557–589.

[28] Luca Pappalardo, Dino Pedreschi, Zbigniew Smoreda, and Fosca Giannotti. 2015a. Using big data to study the link between human mobility and socio-economic development. In *Big Data (Big Data), 2015 IEEE International Conference on*. IEEE, 871–878.

[29] Luca Pappalardo, Filippo Simini, Salvatore Rinzivillo, Dino Pedreschi, Fosca Giannotti, and Albert-László Barabási. 2015b. Returners and explorers dichotomy in human mobility. *Nature communications* 6 (2015).

[30] Peter Pirolli and Stuart Card. 1997. The evolutionary ecology of information foraging. *Palo Alto Research Center Technical Report UIR-R97-01, Palo Alto, CA* (1997).

[31] Peter Pirolli and Stuart Card. 1999. Information foraging. *Psychological review* 106, 4 (1999), 643.

[32] Philip M Podsakoff, Scott B MacKenzie, Jeong-Yeon Lee, and Nathan P Podsakoff. 2003. Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of applied psychology* 88, 5 (2003), 879.

[33] Chirag Shah and Roberto González-Ibáñez. 2011. Evaluating the synergic effect of collaboration in information seeking. In *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*. ACM, 913–922.

[34] Claude Elwood Shannon. 2001. A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review* 5, 1 (2001), 3–55.

[35] Andrew Sih, Alison Bell, and J Chadwick Johnson. 2004. Behavioral syndromes: an ecological and evolutionary overview. *Trends in ecology & evolution* 19, 7 (2004), 372–378.

[36] Rahul Singh, Ya-Wen Hsu, and Naureen Moon. 2013b. Multiple perspective interactive search: a paradigm for exploratory search and information retrieval on the web. *Multimedia tools and applications* 62, 2 (2013), 507–543.

[37] Vivek Kumar Singh, Burcin Bozkaya, and Alex Pentland. 2015. Money Walks: Implicit Mobility Behavior and Financial Well-Being. *PloS one* 10, 8 (2015), e0136628.

[38] Vivek K. Singh, Laura Freeman, Bruno Lepri, and Alex Pentland. 2013a. Predicting spending behavior using socio-mobile features. *Proceedings - SocialCom/PASSAT/BigData/EconCom/BioMedCom 2013* (2013), 174–179.

[39] James NM Smith and Hugh PA Sweatman. 1974. Food-searching behavior of titmice in patchy environments. *Ecology* (1974), 1216–1232.

[40] Chaoming Song, Zehui Qu, Nicholas Blumm, and Albert-László Barabási. 2010. Limits of predictability in human mobility. *Science* 327, 5968 (2010), 1018–1021.

[41] Arkadiusz Stopczynski, Vedran Sekara, Piotr Sapiezynski, Andrea Cuttone, Mette My Madsen, Jakob Eg Larsen, and Sune Lehmann. 2014. Measuring large-scale social networks with high resolution. *PLoS ONE* 9, 4 (2014).

[42] Aaron Striegel, Shu Liu, Lei Meng, Christian Poellabauer, David Hachen, and Omar Lizardo. 2013. Lessons Learned from the NetSense Smartphone Study. *Proceedings of the 5th ACM workshop on HotPlanet - HotPlanet '13* (2013), 51–56.

[43] Michael Tidwell and Patricia Sias. 2005. Personality and Information Seeking Understanding How Traits Influence Information-Seeking Behaviors. *Journal of Business Communication* 42, 1 (2005), 51–77.

[44] Rui Wang, Gabriella Harari, Peilin Hao, Xia Zhou, and Andrew T Campbell. 2015. SmartGPA: how smartphones can assess and predict academic performance of college students. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 295–306.

[45] Ryen W White and Resa A Roth. 2009. Exploratory search: beyond the query-response paradigm (Synthesis lectures on information concepts, retrieval & services). *Morgan and Claypool Publishers* 3 (2009).

[46] David Sloan Wilson, Anne B Clark, Kristine Coleman, and Ted Dearstyne. 1994. Shyness and boldness in humans and other animals. *Trends in Ecology & Evolution* 9, 11 (1994), 442–446.

[47] Kazuo Yano, Sonja Lyubomirsky, and Joseph Chancellor. 2012. Sensing happiness. *IEEE Spectrum* 49, 12 (2012), 32–37.