

Describing Trail Cultures through Studying Trail Stakeholders and Analyzing their Tweets

Abigail Bartolome

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science
in
Computer Science and Applications

Edward A. Fox, Chair

D. Scott McCrickard

Jeffrey Marion

June 8, 2018

Blacksburg, Virginia

Keywords: Hikers, Natural language processing, Topic analysis, Twitter, Technology on the
trail

Copyright 2018, Abigail Bartolome

Describing Trail Cultures through Studying Trail Stakeholders and Analyzing their Tweets

Abigail Bartolome

(ABSTRACT)

While many people enjoy hiking as a weekend activity, to many outdoor enthusiasts there is a hiking culture with which they feel affiliated. However, the way that these cultures interact with each other is still unclear. Exploring these different cultures and understanding how they relate to each other can help in engaging stakeholders of the trail. This is an important step toward finding ways to encourage environmentally friendly outdoor recreation practices and developing hiker-approved (and environmentally conscious) technologies to use on the trail.

We explored these cultures by analyzing an extensive collection of tweets (over 1.5 million). We used topic modeling to identify the topics described by the communities of Triple Crown trails. We labeled training data for a classifier that identifies tweets relating to depreciative behaviors on the trail. Then, we compared the distribution of tweets across various depreciative trail behaviors to those of corresponding blog posts in order to see how tweets reflected cultures in comparison with blog posts. To harness metadata beyond the text of the tweets, we experimented with visualization techniques. We combined those efforts with ethnographic studies of hikers and conservancy organizations to produce this exploration of trail cultures.

In this thesis, we show that through the use of natural language processing, we can identify cultural differences between trail communities. We identify the most significantly discussed forms of trail depreciation, which is helpful to conservation organizations so that they can

more appropriately share which Leave No Trace practices hikers should place extra effort into practicing.

Describing Trail Cultures through Studying Trail Stakeholders and Analyzing their Tweets

Abigail Bartolome

(GENERAL AUDIENCE ABSTRACT)

In a memoir of her hike on the Pacific Crest Trail, *Wild*, Cheryl Strayed said to a reporter in an amused tone, “I’m not a hobo, I’m a long-distance hiker”. While many people enjoy hiking as a weekend activity, to many outdoor enthusiasts there is a hiking culture with which they feel affiliated. There are cultures of trail conservation, and cultures of trail deprecation. There are cultures of long-distance hiking, and there are cultures of day hiking and weekend warrior hiking. There are also cultures across different hiking trails—where the hikers of one trail have different sets of values and behaviors than for another trail. However, the way that these cultures interact with each other is still unclear. Exploring these different cultures and understanding how they relate to each other can help in engaging stakeholders of the trail. This is an important step toward finding ways to encourage environmentally friendly outdoor recreation practices and developing hiker-approved (and environmentally conscious) technologies to use on the trail.

We decided to explore these cultures by analyzing an extensive collection of tweets (over 1.5 million). We combined those efforts with ethnographic style studies of conservancy organizations and avid hikers to produce this exploration of trail cultures.

Dedication

To my mother—my best friend and role model, thank you for your unwavering love and support. To my father—my hiking partner and partner in crime, thank you for always rooting for me and for your words of affirmation. To my grandfather who has always emphasized the value of academics and to my grandmother who constantly reminds me to enjoy life. To Gurkha, for always bringing me joy in my times of stress. And to my boyfriend, for always being on my team.

I owe this thesis to y'all!

Acknowledgments

I would like to thank those who have made the completion of this project possible:

- My advisor and committee chair, Dr. Edward A. Fox. Thank you for your constant support, your sage guidance, and for giving me the freedom to study a subject that I am very passionate about.
- My committee member and Technology on the Trail professor, Dr. Scott McCrickard. Thank you for exposing me to the intersection of Human-Computer Interaction and trail research, and for encouraging me to present my research at the Technology on the Trail and HCI Outdoors workshops.
- My committee member and trail expert, Dr. Jeffrey Marion. Thank you for providing me with countless resources on conservation, and for inviting me to chaperon on select Venture Out outings.
- The National Science Foundation for funding the Integrated Digital Event Archiving and Library (IDEAL) (Grant IIS-1319578) and Global Event and Trend Archive Research (GETAR) (Grant IIS-1619028) projects.
- The Digital Library Research Laboratory—especially, Matthew Bock for his framework that produced my topic analysis, and Liuqing Li for helping me collect tweets.
- Shuo Niu for listening to me brainstorm examples of trail cultures and for helping me build a classifier.
- The CS 4624 (Multimedia, Hypertext, and Information Access), CS 4784 (Technology on the Trail), CS 5604 (Information Retrieval), CS 6604 (Digital Libraries) classes.

Contents

List of Figures	x
List of Tables	xii
1 Introduction	1
1.1 Background	1
1.2 Problem	1
1.3 Motivation	3
1.4 Research Questions	5
1.4.1 RQ1: What are the cultural differences among the Triple Crown trails?	5
1.4.2 RQ2: What specific depreciative behaviors do hikers tweet about more than others?	5
1.5 Organization of Thesis	5
2 Related Work	7
2.1 Trail Literature	7
2.2 NLP Literature	9
2.3 Online Community Literature	11
3 Methods	13

4	Preliminary Studies	16
4.1	Triple Crown Trails	16
4.2	Backpacking Cultures	18
4.3	Conservation Culture	22
5	Experiments	31
5.1	Experiment to Describe the Triple Crown Trails	31
5.2	Experiment to Explore Depreciative Behavior	37
5.3	Validating Depreciative Behaviors by Exploring Blog Posts	46
5.4	Visualizing Tweets on the Trail	48
6	Conclusion	54
6.1	Results	54
6.1.1	RQ1: What are the cultural differences among the Triple Crown Trails?	54
6.1.2	RQ2: What specific depreciative behaviors do hikers tweet about more than others?	56
6.2	Discussion	56
6.3	Limitations	58
6.4	Future Work	60
7	Summary	64
	Bibliography	65

Appendices	72
Appendix A Map of Triple Crown Trails	73
Appendix B ACM GROUP 2018 Technology on the Trail Workshop	75
Appendix C ACM CHI 2018 HCI Outdoors Workshop	80

List of Figures

1.1	“Heavy horse traffic has compacted and incised the main tread, which captures and retains water; subsequent hikers and horse riders seeking to avoid mudholes widen trails. Leave No Trace guidelines ask visitors to stay as close to the center of the tread as possible to avoid trail widening.”—Marion et al. [33]	2
3.1	Workflow for approaching the problem	14
3.2	Outline of qualitative and quantitative work	15
4.1	Jeffrey Marion on a reunion thru-hike on the Appalachian Trail [32]	18
4.2	Comparing two 2-person backpacking tents by MSR on REI.com	20
4.3	Tent Reviews for the Tents in Figure 4.2 from REI.com Top Review—the lighter tent helped with the need to reduce weight Bottom Review—the heavier tent was a good tent, but is heavier, which is a con to the purchaser	21
4.4	Three principles of Leave No Trace are: plan ahead and prepare, travel and camp on durable surfaces, dispose of waste properly. These principles are described in this hangtag, with respect to the Appalachian Trail.	25
4.5	Four principles of Leave No Trace are: leave what you find, minimize campfire impacts, respect wildlife, and be considerate of other visitors. These principles are described in this hangtag, with respect to the Appalachian Trail.	26
4.6	Options for starting an Appalachian Trail thru-hike [15]	28

5.1	Circle packing visualization of topic analysis on Appalachian Trail tweets . . .	36
5.2	Circle packing visualization of topic analysis on Pacific Crest Trail tweets . . .	37
5.3	TreeMap visualization of topic analysis on Appalachian Trail tweets	38
5.4	TreeMap visualization of topic analysis on Pacific Crest Trail tweets	39
5.5	Example of good behavior on a hike to counteract trail depreciation	44
5.6	Example advocating against trail depreciative behaviors	44
5.7	Example of showcasing trail depreciative behavior	45
5.8	Word cloud on hiking tweets [21]	49
5.9	Heatmap of tweets in Asia [21]	50
5.10	Heatmap of tweets in Europe [21]	51
5.11	Map of tweets from the United States with a layer turned on to mark the season during which it was tweeted [21]	52
5.12	Map of tweets from the United States with a filter that only displays tweets containing the search term “dog” [21]	53
6.1	Heatmap of tweets in North America [21]	63
A.1	Map of Triple Crown Hikes [27]	74

List of Tables

4.1	Mission statements of conservancy organizations and their respective trails	17
4.2	Training for backpacking with minimum backpacking—to be supplemented with a wholesome diet	23
5.1	Details of tweet collection	33
5.2	Results of topic analysis of tweets: January-May	35
5.3	Details of tweet collection—one year of collecting	41
5.4	Training data—breakdown of tweets for depreciative behaviors	42
5.5	Depreciative behavior categories with more than 10 tweets	43
5.6	Manually evaluating the performance of the classifier	46
5.7	Blogs mentioning depreciated behavior	48

Nomenclature

AT is an abbreviation for Appalachian Trail. The AT is a national scenic trail that runs from Georgia to Maine.

Backpacker is the term for a hiker who carries multiple days of hiking and camping gear in their backpack for a multi-day hike.

CDT is an abbreviation for Continental Divide Trail. The CDT is a national scenic trail that runs from the Mexico/New Mexico border to the Montana/Canada border.

LDA is an abbreviation for Latent Dirichlet Allocation. LDA is a natural language processing technique to find topic distributions over a corpus of text.

NLP is an abbreviation for Natural Language Processing. NLP is a field of computer science, artificial intelligence, and linguistics concerned with the interactions between computers and human (natural) languages.

PCT is an abbreviation for Pacific Crest Trail. The PCT is a national scenic trail that runs from the Mexico/California border to the Washington/Canada border.

Section hiker is the term for a backpacker who is hiking a section of a long distance trail over multiple days at a time.

Thru-hiker is the term for a backpacker who is spending months hiking the entirety of a long distance trail. Many groups will define a hike as a thru-hike if the hiker completed the full trail in 12 months.

Trail journals are written by backpackers to reflect on their hikes. The website, www.trailjournals.com,

is a website that hosts a community of hikers and provides them a place to keep their journals and where the public can read about their hiking adventures.

Chapter 1

Introduction

1.1 Background

Hiking is an activity that can be enjoyed by many. For many people, from tourists wanting to see the iconic Half Dome in Yosemite National Park to thru-hikers standing atop Mount Katahdin after completing a northbound Appalachian Trail thru-hike (i.e., thru-hiking the full length of a long distance trail in one hiking season), a vigorous walk on a trail is a great way to enjoy nature. Because hiking is an activity that is so versatile and accessible to all, it is difficult to make a generalization of hikers—what they value, what they look for in trails, how they respect nature, etc.

1.2 Problem

For trail specialists, this can be a huge obstacle in their work. Each year, thousands of hikers attempt to thru-hike the Appalachian Trail. Only hundreds of hikers attempt to thru-hike the Pacific Crest Trail, the West Coast sibling of the Appalachian Trail [7]. Furthermore, most of the thru-hikers who attempt the Appalachian Trail are hiking northbound (from Georgia to Maine) [13]. With trail usage as high as it is on the Appalachian Trail, its resources are continuously strained and its need for efforts in conservation is escalating [16]. While this need is evident to experts who study trails, it is not necessarily at the forefront of every

recreational hiker’s mind. Forestry scientists and conservancy organizations have to find ways to bring the public’s attention to the importance of conservation, along with ways to aid in the effort. Figure 1.1 shows an example of hikers who are hiking on a trail, completely unaware that they are contributing to widening the trail by trying to circumvent the muddy areas of the trail. By exploring communities and cultures on the trail, we can better understand the values and motivations of recreational hikers and so help trail specialists in influencing hikers to adapt to hiking practices that are in line with the effort to conserve trail resources on the Appalachian Trail, as well as on other major hiking trails.



Figure 1.1: “Heavy horse traffic has compacted and incised the main tread, which captures and retains water; subsequent hikers and horse riders seeking to avoid mudholes widen trails. Leave No Trace guidelines ask visitors to stay as close to the center of the tread as possible to avoid trail widening.”—Marion et al. [33]

1.3 Motivation

This work was inspired by the author's enthusiasm for hiking and the great outdoors. Through her experiences on various hiking trails in different parts of the country, she has become aware of their cultural differences. During a section hike on the Appalachian Trail in Great Smoky Mountain National Park, she spoke with a ridge runner who likened the Appalachian Trail to a religion. This ridge runner supported this conjecture by citing the mission statement and vision of the Appalachian Trail Conservancy, as well as testimonials of its thru-hikers, and stating that the language in these writings were far more spiritual than those of the Pacific Crest Trail and Continental Divide Trail. The following excerpt comes from the author's position paper for the Technology on the Trail workshop at ACM GROUP 2018 [5]. The full paper can be found in Appendix B.

This casual conversation on my backpacking trip was incredibly impactful, and it inspired my research efforts for the following two years. This ridge runner had a point. Not all trails are the same, including the Triple Crown Trails. What is it about these trails that makes them each so different? The Appalachian Trail, Pacific Crest Trail, and Continental Divide Trail are all national scenic trails whose lengths are well over 2,000 miles. Needless to say, hikers that embark on hiking these trails in one straight push must be particularly ambitious and have some sort of connection to the great outdoors. But there must be some difference between hikers that decide to take on the Appalachian Trail versus hikers that take on the Pacific Crest Trail—it cannot just be a matter of geographical convenience, right? And even if the trail that a hiker chooses is based on geographical convenience, hikers on each of the respective trails must have very different experiences. Consider Dr. Ellie Harmon's source of maps on her Appalachian Trail thru-hike versus her Pacific Crest Trail hike. When she hiked the Appalachian

Trail, like many, all she needed was a guidebook as her source of direction. However, when she hiked the Pacific Crest Trail, she required her phone as her guide. She wrote in her dissertation that her phone was her everything on her Pacific Crest Trail thru-hike [24]. I wanted to dedicate my research to learning what makes each of these trails different.

After conducting more research on trail cultures, it became evident that the study should go beyond a mere comparison among the trails. For example, across all trails, there are subcultures that can be classified by the intensity of the hikers. The vernacular and practices of long distance hiking cultures are different from that of a hiking tourist. We posit that this study should be done with natural language processing on tweets from the hiking community because of the informal nature of Twitter. Twitter has a 280 character limit (140 character limit before the fall of 2017). This drives the need to keep tweets brief, so many Twitter users feel enabled to tweet real-time updates. The Digital Library Research Laboratory has collected over 1.5 million tweets related to hikes to make this research possible. In this thesis, we explore specific cultures on trails so as to give these cultures meaningful descriptions that will benefit trail researchers and developers of technology on the trail. This thesis will specifically focus on the cultural differences among the Triple Crown trails and the culture of depreciative behavior on trails.

1.4 Research Questions

1.4.1 RQ1: What are the cultural differences among the Triple Crown trails?

What can tweet analysis do to describe trail cultures? What causes a hiker to choose a particular trail?

1.4.2 RQ2: What specific depreciative behaviors do hikers tweet about more than others?

What categorizations of tweets represent the depreciative behaviors on the trail? Are tweets about trail depreciation more focused on advocating against, engaging in, or showing evidence of trail depreciation?

1.5 Organization of Thesis

This thesis is organized as follows. Chapter 2 is a review of related works which discusses our readings about different trails and trail cultures, as well as literature which describes the natural language processing techniques used in our study. Chapter 3 describes the methods used to conduct the research for this thesis. Chapter 4 describes the preliminary studies that were used to provide an in-depth description of trail cultures. Chapter 5 describes the experiments that were conducted to provide quantitative data for this work. Chapter 6 presents our conclusions, limitations, and proposals for future work. Chapter 7 summarizes the work of this thesis. This thesis contains excerpts from position papers written by the

author in an effort to further this work. These position papers can be found in [Appendix B](#) and [Appendix C](#).

Chapter 2

Related Work

2.1 Trail Literature

In *America's Greatest Hiking Trails*, Triple Crown veteran Karen Berger discussed each trail in a way that indicated they each had a particular identity. She assigned the Appalachian Trail the moniker of “A Community in the Wilderness” for its ability to attract hikers from a vast assortment of backgrounds, and for its remarkably large number of annual thru-hike attempts. She describes the Continental Divide Trail as “The Wild Child”, since those attempting to thru-hike the incomplete trail must often “make their own route, following the suggestions of those who have gone before, or improvising their own paths through trailless wetlands”. She referred to the Pacific Crest Trail as “The Soul of American Wilderness”, because unlike the AT, which prides itself on being a community in the wilderness, and the CDT, which is known for its hikers conquering nature, the PCT is about “melding with the wilderness and sometimes submitting to it”. Since she writes about each trail from her own personal knowledge and experience, the judgments that were made for assigning each trail an identity are personal to her. We suggest that, by aggregating the knowledge and experiences of a larger portion of each trail community, we could better understand that collective community's values, and so, we could better define what the community of that trail values [7].

The Alan Walks Wales project demonstrated the value of this effort, since in Alan Dix's CHI 2013 GeoHCI (Workshop on Geography and HCI) paper, “Mental Geography, Wonky Maps, and a Long Way Ahead”, he writes “allowing open government data and other public (or leaked) data to be visualized in ways that may subvert or offer alternative views to the official narrative”. We are particularly drawn to the phrase “offer alternative views” because that is often what happens when hikers share their experiences on trails. The memories and impressions that hikers have on a trail are all dependent on that person's hiking experience, the weather on the trail, the time of the year, and the other present company on the trail. However, while the recollection of a trail will vary from hiker to hiker, there are still characteristics of a trail that are constant. Those are the characteristics that draw hikers to the trail. Those are the characteristics that give the trail its identity [19].

Norman Su wrote a paper on American Midwest Hunters, which justified this research effort by showcasing different practices in hunting and their diverse approaches to fair chase ethics. “Reconsidering Nature: The Dialectics of Fair Chase Practices of American Midwest Hunters” illustrates an ethnographic study of the practices and dialects of American Midwest hunters. His study shows that within the community of hunters in America's Midwest, there are different sub-communities. Similarly, within the community of hikers in the United States, there are different sub-communities—three of which are those who associate with the AT, PCT, and CDT [44].

Ellie Harmon's “Stories of the Smartphone in everyday discourse: conflict, tension & instability” [24], as well as her dissertation about what it means to connect/disconnect, were particularly influential to my decision to study how each community viewed the use of technology on the trail [23]. The trope of individuals who view non-users [of smartphones] to be

an authentic human reflected what we would have expected of a trail community that was against using technology on the trail. Whereas, the trope of individuals who supported using [smartphones] by the promises of integration reflected the hikers that we suspected would advocate for using technology on the trail as a means of safety and making trails more accessible. Ellie Harmon's dissertation was particularly compelling because she relied so heavily on her smartphone for her PCT thru-hike. She used her smartphone as her map, and she used her smartphone for music when she felt overwhelmed by the monotony of hiking all day. She literally titled her dissertation chapter on hiking the PCT, “My Maps, My Music, My Everything”.

During a Technology on the Trail workshop session at ACM GROUP 2018, led by Professor Michael Horning in Virginia Tech's Department of Communication, and Lindah Kotut, Ph.D. student in Virginia Tech's Department of Computer Science, participants clustered various stakeholders of technology on the trail (e.g., thru-hikers, section hikers, day hikers, trail angels, ridge runners, park rangers, farmers, families, etc.), and examined how they might benefit from developing technology on the trail. During these discussions, participants discussed the values and cultures of these clusters and so allowed us to identify different trail cultures.

2.2 NLP Literature

There are many studies that have used natural language processing to study tweets. In 2017, Jamil et al. built a series of classifiers to automate the process of find tweets from users that were at-risk of depression [28]. In 2018, Mullick et al. built a generic classifier to identify fact and opinion. They used Multi-perspective Question and Answering articles, Yahoo! news

articles, Twitter hashtag idioms, and YouTube comments to build this classifier [38]. Also in 2018, Li et al. used natural language processing techniques such as Linguistic Inquiry and Word Count (LIWC) and descriptive statistical analysis to understand how immigrant-focused non-profit organizations' used Twitter during times of political conflict [29].

To perform topic analysis, we had to define what we were looking for in a topic. We defined a topic as the main idea discussed in a document, and we represented the topic as a distribution of words [3]. In Chapter 17 of *Text Data Management and Analysis*, topic analysis was divided into two distinct processes: discovering topics and seeing which documents belong to those topics. Making a probability distribution over the words allows the topics to be represented as a word distribution that allows topics to be expressive and complex. This can be done with Latent Dirichlet Allocation, an unsupervised learning method that allows us to represent the topics as a probability distribution [45].

Latent Dirichlet Allocation (LDA) is a popular topic modeling approach connected with natural language processing. LDA represents a topic as a word distribution; that allows the topics to be more expressive and deal with more complicated issues, by making a probability distribution over the words [45]. LDA is a probabilistic model of a corpus, and treats the documents as random mixtures over latent topics. LDA uses a three-level model with a topic node that is repeatedly sampled, thus allowing documents to be associated with multiple topics [8].

The NLP efforts of this project were driven by the efforts of Virginia Tech's CS 5604 class. The class is conducted with a project-based learning approach that allowed the team to focus on clustering the more than 1.5 billion tweets collected by the Digital Library Research Lab-

oratory (DLRL). In the Spring 2016 semester, the topic analysis team used Apache Spark’s MLlib library to build topic models from tweets that had already been preprocessed (removed profanity and stopwords and lemmatized text), using LDA. Each tweet acted as a document, and LDA produced a topic distribution for each of the documents, where each topic was a distribution of related terms. From the topic distribution, the team used a document matrix to determine which topic each document belonged to [37].

From the efforts of the CS 5604 class, Matthew Bock, M.S., was able to include an `LDARapper` tool in the framework that he built as his thesis work [9]. Since `LDARapper` allowed me to refine various LDA parameters (i.e., number of topics, number of iterations, terms to ignore), we were able to leverage the tool and extract meaningful topics that the trail communities were discussing over Twitter.

2.3 Online Community Literature

The idea that we could study the values of trail communities stemmed from Danescu-Niculescu-Mizil et al. In *No Country for Old Members*, they noted that members of an online beer community “helped define the collective identity of the community as it stands today” [18]. They showed that linguistic change in online communities was a result of evolving social norms, and consequently, evolving values.

Stevie Chancellor, Andrea Hu, and Munmun De Choudhury wrote a remarkable paper that notes the social norms in online health communities. They determined that there were two forms of support in online health communities: encouraging healthy behavior change or encouraging harmful behavior change. To understand the different types of encouragement in

online health communities, they studied the norms of the two communities to identify the differences in how they encourage behavior change. Their study showed that the norms of a community are influential in supporting either type of behavioral change [10].

The interest in using tweets to understand trail communities was affirmed in “Imagining Twitter as an Imagined Community”. In 2011, Gruzdt et al. stated that “Twitter can form the basis of interlinked personal communities—and even of a sense of community” [22]. In the study, a Twitter community was “imagined”. The community had members who interacted and felt a sense of interpersonal commitment. The study found that the imagined community was helpful in “developing theories of online community building”. This success was attributed to having a group of members that were integral to community in their participation, both in actively interacting with each other and in being part of the community for a long time. Its success was also attributed to its openness to newcomers. Since a Twitter network is asymmetrical (i.e., “You can follow me but I do not have to follow you”) [22], it allows new members to connect to the network and assimilate to the norms of a network and thus form a community.

In 2009, Dejin Zhao and Mary Beth Rosson conducted an exploratory study to understand why people tweet. This study was conducted in the early stages of Twitter (less than two years after the site was launched) and aimed to bring an understanding of why people use microblogs, such as Twitter, into an academic context. Through conducting interviews, they found that the desire to use Twitter was driven by the desire to learn about others' thoughts and personal updates, build common ground between their network connections, and consequently grow in connectedness and “sustain a virtual feeling of proximity” [46].

Chapter 3

Methods

Our problem solving approach is depicted in Figure 3.1. We began with a study using qualitative analysis that explored the trail communities and their cultures. In order to accurately describe trail cultures, it was imperative that we start by first studying the trails and their stakeholders. We did preliminary studies by researching the mission statements of the Triple Crown trails and reading the reflections of Karen Berger who has thru-hiked all three Triple Crown Trails. We then looked at different subcultures within general trail culture, focusing on backpacking culture and conservation culture, which have unique vocabularies and values associated with them. By conducting preliminary studies, which allowed us to intimately understand the needs of the trails and their communities, we were able to better prepare to analyze their tweets because we understood their needs and values.

We validated our findings from the qualitative analysis by conducting experiments to find the unique topics in the tweets relating to each of the triple crown trails. We used a natural language processing technique called Latent Dirichlet allocation to extract the topics from these tweets. These topics were then validated by using a tool developed by Virginia Tech's Spring 2018 CS4624 (Multimedia, Hypertext, and Information Access) class's Trail Study team, which allowed us to search for those topics in a corpus of blog posts that were written by hikers on the Triple Crown trails. We also used classification, a natural language processing technique that uses training data to classify new datasets into the predetermined categories,



Figure 3.1: Workflow for approaching the problem

to identify various depreciative behaviors in tweets. The process of labeling the training data was particularly useful because it enabled us to see if there were people who tweeted about participating in trail depreciation or if the tweets relating to trail depreciation were by those who advocated against it.

This methodology allowed us to understand the values and motivations of recreational hikers. We intend for this work to help trail specialists influence hikers to adapt conservation conscious hiking practices. The qualitative and quantitative studies can be better visualized in Figure 3.2.

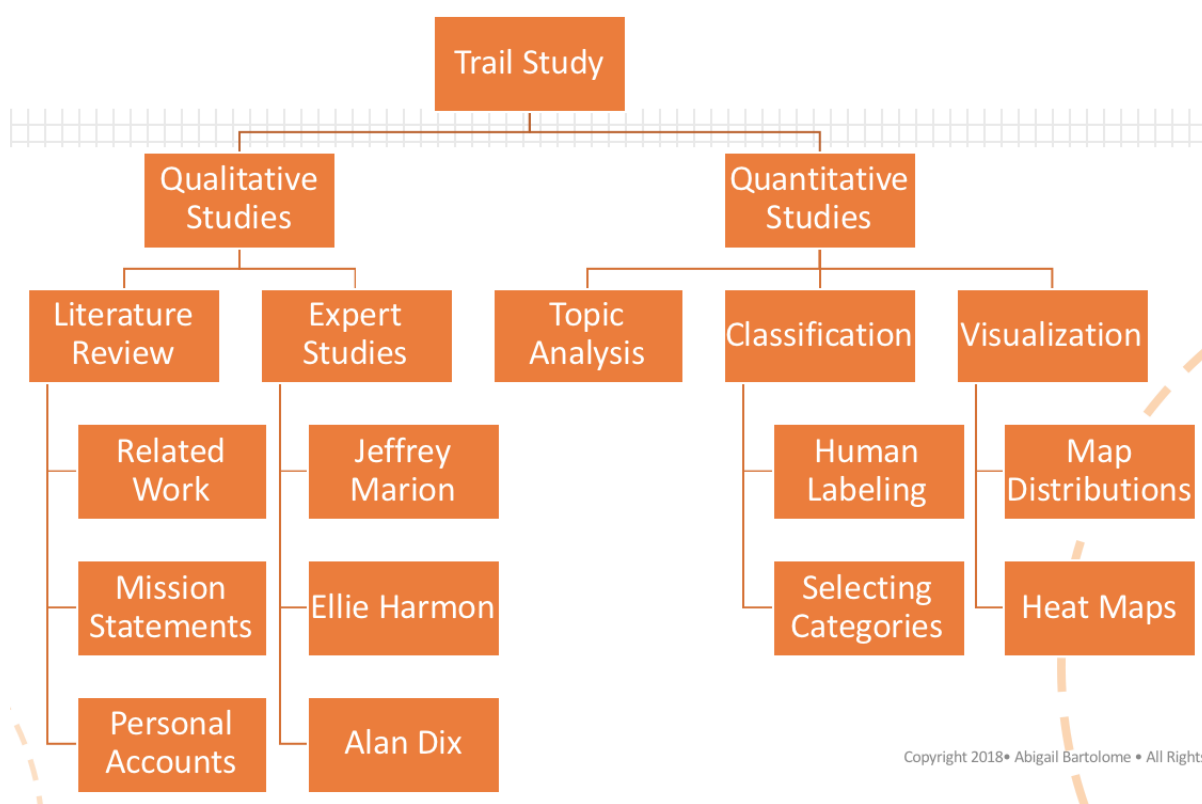


Figure 3.2: Outline of qualitative and quantitative work

Chapter 4

Preliminary Studies

4.1 Triple Crown Trails

To have an objective understanding of the Triple Crown trails, it was important to conduct preliminary studies on each of the trails. We studied the mission statements of each of the Triple Crown conservancy organizations to see if an identity could be identified for each trail. Table 4.1 shows these mission statements. From the mission statements, it appeared that the Appalachian Trail's identity was focused on preserving “natural beauty and priceless cultural heritage”. The Pacific Crest Trail Association’s mission statement reflects an identity that values a “world-class experience” for those who use that trail. Meanwhile, the Continental Divide Trail showed that it values “power and grandeur”. At a Technology on the Trail workshop at Virginia Tech, Professor Alan Dix of the School of Computer Science at the University of Birmingham commented on the mission statements, stating that they were written by public relations and marketing experts who use “flowery” and generalized language to describe the trails. To put it succinctly, the mission statements were too innocuous to truly embody the spirit of the trails, because their mission statements were written in such a way to attract people to explore and protect their respective trails. This further motivated the effort to study the tweets of the Triple Crown trails in order to find each trail's unique identity.

Karen Berger, a Triple Crown veteran, found it difficult to pick a favorite of the three trails.

Conservancy Organization	Trail	Mission Statement
Appalachian Trail Conservancy	Appalachian Trail	The Appalachian Trail Conservancys mission is to preserve and manage the Appalachian Trail – ensuring that its vast natural beauty and priceless cultural heritage can be shared and enjoyed today, tomorrow, and for centuries to come. [14]
Continental Divide Trail Coalition	Continental Divide Trial	Inspired by the power and grandeur of the CDT and in keeping with the Trail’s values, CDTC commits to conducting all transactions and dealings with integrity and honesty, and promoting working relationships with board members, staff, volunteers, partners and program beneficiaries that are based on mutual respect, fairness, and openness. [12]
Pacific Crest Trail Association	Pacific Crest Trail	The mission of the Pacific Crest Trail Association is to protect, preserve and promote the Pacific Crest National Scenic Trail as a world-class experience for hikers and equestrians, and for all the values provided by wild and scenic lands . [2]

Table 4.1: Mission statements of conservancy organizations and their respective trails

As she described the three trails in her *SectionHiker.com* essay, it was evident that their differences were beyond the physical traits (e.g., trail mileage, geographic location, wildlife, etc.). Her exact words were, “each one is different; each has a soul” [6].

4.2 Backpacking Cultures



Figure 4.1: Jeffrey Marion on a reunion thru-hike on the Appalachian Trail [32]

In a position paper to the HCI Outdoors workshop at ACM CHI 2018, we described backpacking as a trail culture [4]. The full paper can be found in Appendix C.

When many people imagine the stereotypical hiker, they may picture a hiker with days worth of gear strapped to their back, such as the hikers depicted in Figure 4.1.

These hikers are called backpackers. Some backpackers may go out for a few

days at a time (section hikers), while the more devoted may go out for several months at a time to complete a thru-hike. Whether a backpacker is out for a 50 mile hike over a 3 day period or spending months thru-hiking all 2,190 miles of the Appalachian Trail, backpackers are a devoted group of hikers with their own unique culture.

In a reflective article comparing modern backpacking to backpacking in the 1970's, Jeffrey Marion writes, "No backpacker with a few thousand miles under his feet fails to consider the weight of every packed item. The ultra-light paradigm places greater focus on a new array of light-weight gear, your knowledge and skills, and doing without" [32]. Managing trail weight is a common culture in backpacking. There are extreme practices like shortening the handle of a toothbrush or repackaging food to lessen the weight of the backpack. Such practices seem extreme to those outside of the backpacking community. The term "comfort item" is often used by backpackers to refer to items considered "non-essential," that make the backpack heavier than it needs to be. However, when backpackers discuss bringing a 1-pound camping chair or a warmer, fluffier (i.e., heavier) sleeping bag because it makes their trail experience more enjoyable, the untrained ear may struggle to understand how the difference between a 29-pound backpack and a 30-pound backpack could cause such a disturbance to a physically fit backpacker. In the backpacking community, the notion of spending an extra \$100 on a smaller tent is justifiable if it means reducing trail weight. Figure 4.2 shows a comparison of two backpacking tents. Both are 2-person, 3-season, and made by MSR. The main difference between the Hubba Hubba and the Elixir is that the Hubba Hubba weighs about half as much as the Elixir, and accordingly costs an extra \$150. Oftentimes, in backpacking culture, saving the 2+ pounds in trail weight

is well worth the \$150.

The purpose of this study was to better shed light on the backpacking community, a strong community of stakeholders on the trail. We wanted to showcase how this community has niche values that others may not consider or value. For example, Figure 4.3 showcases how trail weight is so valuable to hikers that it is a factor in their tent reviews. The community speaks in a language that may not make sense to those outside of the community. “Trail weight”, “comfort item”, and even “thru-hike” belong to a vocabulary that not many people outside of the backpacking community would use.



		
	MSR Hubba Hubba NX 2-Person Tent	MSR Elixir 2 Tent With
	\$399.95	\$249.95
	Shop item	Shop item
	Remove	Remove
	Backpacking	Backpacking
	3-season	3-season
Capacity	2-person	2-person
Trail Weight	3 lbs. 7 oz.	5 pounds
l Weight	3 lbs. 13 oz.	6 pounds

Figure 4.2: Comparing two 2-person backpacking tents by MSR on REI.com

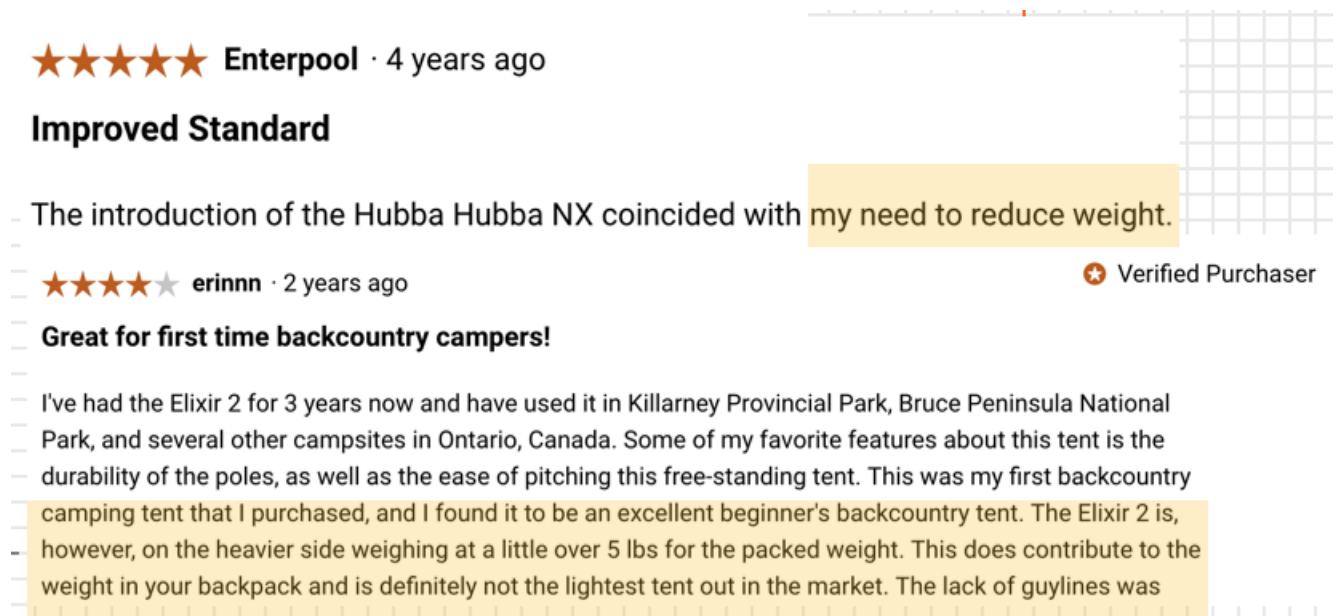


Figure 4.3: Tent Reviews for the Tents in Figure 4.2 from REI.com

Top Review—the lighter tent helped with the need to reduce weight

Bottom Review—the heavier tent was a good tent, but is heavier, which is a con to the purchaser

We extended this study with an exercise in autobiographical design [25]. Neustaedter and Sengers define autobiographical design as “design research drawing on extensive, genuine usage by those creating or building the system” [40]. In this exercise, we were stakeholders in the design of a training program for backpacking, that included minimal backpacking. The author of this thesis is a self-described hiking enthusiast with an affinity for hiking, backpacking, camping, and reading about gear. However, she does not always enjoy getting in shape for a major backpacking trip. On an Appalachian Trail section-hike, a thru-hiker told her “the only way to get in shape for backpacking is to actually go backpacking”. However, for many, it is difficult to set aside the responsibilities of everyday life in order to go on weekend backpacking trips. Unsatisfied by the results of the ambiguous training regimen of “running and eating healthy”, the author of this paper designed the training plan in Table 4.2. Jaya Marr, from the Mountain Training School in Chile had a “Be as Fit as a Mountain Guide”

fitness plan to get in mountaineering shape in just 12 weeks [34]. However, the workouts favored strength training over aerobic exercise, which did not cater to the needs of a hiker who was just going on a 5-day backpacking trip. While the program was appealing in its quick and effective means for getting “mountain ready” in 12 weeks, backpacking requires less of a focus on muscular strength and more of a focus on building a stronger stamina. The author tailored the Mountain Training School's program by taking the valuable pieces of the program (i.e., starting at a slow pace of running to focus on endurance building), and she replaced some of the less essential strength exercises with more cardio. The author also added day hikes as a celebration of sorts to appreciate the hard work that has been put into training, as well as to acclimate the body to the weight of a backpack and the changing elevations that commonly occur on hikes.

After designing this training program, the author presented it to Virginia Tech's Fall 2017 Research through Design class [25], as well as her father (backpacking partner). There was feedback to incorporate StairMaster exercises into the training program to prepare for steeper ascents, and that feedback will be considered for future iterations of the design process. Ultimately, this design exercise was valuable in understanding the values and training rituals of backpacking culture.

4.3 Conservation Culture

Also in a position paper to the HCI Outdoors workshop at ACM CHI 2018, we discussed conservation as a trail culture [4]. The full paper can be found in Appendix C.

The outdoor spaces that are so loved and coveted by trail enthusiasts are protected by hikers and conservancy organizations. One well-known conservancy

Week	Exercise Plan
Week 1	Run for 45 minutes 4 days a week 2 days of core training
Week 2	Run for 60 minutes 4 days a week 2 days of core training
Week 3	Run for 60 minutes 4 days a week 2 days of core training 1 long run—1.5 hours minimum
Week 4	Run for 60 minutes 4 days a week 2 days of core training 1 long run—1.5 hours minimum
Week 5	Run for 60 minutes 4 days a week 2 days of core training 1 long run—1.5 hours minimum
Week 6	Run for 60 minutes 3 days a week 1 day of core training 1 long run—1.5 hours minimum 1 hike with a 15 pound pack—10 miles minimum
Week 7	Run for 60 minutes 3 days a week 1 day of core training 1 long run—2 hours minimum 1 hike with a 20 pound pack—8 miles minimum
Week 8	Run for 60 minutes 3 days a week 1 day of core training 1 hike with a 25 pound pack—12 miles minimum
Week 9	Run for 60 minutes 3 days a week Maintain fitness
Week 10	Run for 60 minutes 3 days a week Maintain fitness

Table 4.2: Training for backpacking with minimum backpacking—to be supplemented with a wholesome diet

organization is the Leave No Trace Center for Outdoor Ethics. Practicing Leave No Trace is common in the conservation culture. Practices as common as “pack-it in, pack-it out” are part of Leave No Trace. Other less visible practices such as digging cat-holes (holes for solid human waste that are dug 4-6 inches in diameter and 6-8 inches in depth) 200 feet from the trail and water sources, or minimizing visible impact after camping in pristine areas, also are part of Leave No Trace, and thus conservation culture [30].

The purpose of this study was to highlight a culture of stakeholders who love and respect the trail so much that they want to take care of it. This culture is so passionate about caring for the outdoors that they engage in such careful and well thought-out practices that those outside of the zeitgeist may not even consider, such as cautiously disposing of used toothpaste by spitting toothpaste into rinse water and disposing of it far from a campsite or “broadcasting” their spit over a wide area. Leave No Trace (LNT) is a conservancy program that advocates for practicing low-impact outdoor skills [30]. The principles of Leave No Trace, further described in the Appalachian Trail Leave No Trace hangtag depicted in Figures 4.4 and 4.5, are as follows.

1. Plan Ahead and Prepare
2. Travel and Camp on Durable Surfaces
3. Dispose of Waste Properly
4. Leave What You Find
5. Minimize Campfire Impacts
6. Respect Wildlife
7. Be Considerate of Other Visitors



Outdoor Ethics for the Appalachian National Scenic Trail

Plan Ahead and Prepare

- Pack the clothing and gear you'll need for protection from cold, wind, and rain.
- Do not rely on space in a shelter: Carry a tent or waterproof tarp.
- Bear canisters are required in some areas; elsewhere, use a 50 ft rope and a waterproof bag to hang food via the "PCT method."
- Carry maps and/or reliable guidebook(s)
- Pack a camp stove for cooking. Carry a lightweight trowel or wide tent stake for digging a 6-8" deep hole for burying human and pet feces.

Travel and Camp on Durable Surfaces

- Stay on the trail; don't shortcut switchbacks and cause erosion.
- Follow camping and campfire restrictions that many areas have in place for landscape and watershed protection.
- Use existing campsites out of sight of the Trail and at least 200 feet (80 steps) from water.
- Remove branches from the Trail whenever possible, rather than going around and creating a new path.

Dispose of Waste Properly

- Use a litter bag. Instead of burning any trash or food scraps, pack all of it out.
- Disperse grey water (urine, toothpaste, strained dishwater and cooking water) at least 100 ft (40 steps) away from campsites and shelters and at least 200 ft (80 steps) away from water sources. This protects the soils and plants near campsites, helps keep rodents and other wildlife away from sleeping areas, and preserves water quality.
- Bury fecal waste (from humans and pets) well away from the trail, at least 200 ft from campsites, shelters and water sources, or ***preferably miles from areas where humans concentrate,*** by digging a 6-8" hole. Use a stick to push toilet paper to the bottom of the hole, or else pack it out. Use privies (outhouses) only for fecal waste, not for food or other garbage.

Figure 4.4: Three principles of Leave No Trace are: plan ahead and prepare, travel and camp on durable surfaces, dispose of waste properly. These principles are described in this hangtag, with respect to the Appalachian Trail.

- Keep water sources clean. Wash or rinse clothing, dishes and bodies at least 200 ft from water sources; do not introduce soap, food, dirty dishes, sunscreen or insect repellent into lakes, streams or springs.



Leave What You Find

- Don't mark trees, rocks, structures (shelters, bridges, signs) or historic sites. Leave them as you find them, so future visitors can enjoy them as they are.
- Use rubber tips on trekking poles to avoid leaving scratches on rocks and eroding the soil.
- Good campsites are found, not made. Avoid trampling or removing plants. Don't dig trenches or build structures in your campsite.
- Respect property, public and private. Leave gates as you find them.

Minimize Campfire Impacts

- Use a backpacking stove for cooking. Build a fire only where it's legal, using an approved and existing fire ring. Keep your fire small.
- Leave hatchets and saws at home. Gather small pieces of dead wood from the ground that can be broken by hand.
- Don't burn food waste or other trash.
- Burn all wood to ash and be sure that the fire is completely out and cold, and the fire ring clean, when you depart.

Respect Wildlife

- Observe animals from a distance and never approach, feed or follow them.
- Human food is unhealthy for wildlife, and feeding them causes them to lose their natural fear of humans, leading to conflicts and restrictions such as closed campsites or other regulations.
- Protect wildlife and your food by using approved food storage devices for food, trash and scented articles.
- Control and leash your pets at all times, or leave them at home.
- Wildlife and farm animals are not pets. Do not follow, approach or feed them.

Be Considerate of Other Visitors

- Think about whether your fun is intruding on others. Respect nature's quiet. Avoid yelling.
- Don't "take over" a shelter; be welcoming, and limit how long you stay.
- In most places, groups over 5 should camp, rather than staying in shelters.
- Step off on a durable surface (for example, a rock) and yield to others on the trail, especially hikers going uphill.
- Respect the people who live and work along the Trail.

For more information:

Videos: <https://goo.gl/yJEffC> "Don't Be That Guy on the Appalachian Trail"

Facebook: "Appalachian Trail Leave No Trace" • Social media: #ATLNT

Appalachian Trail Conservancy: www.appalachiantrail.org/LNT

Leave No Trace Center for Outdoor Ethics: www.LNT.org 1.800.332.4100

Figure 4.5: Four principles of Leave No Trace are: leave what you find, minimize campfire impacts, respect wildlife, and be considerate of other visitors. These principles are described in this hangtag, with respect to the Appalachian Trail.

This care for hiking trails extends all the way to planning the logistics of a hike. The Appalachian Trail Conservancy (ATC) has been concerned that the rise of hikers attempting to thru-hike the AT has led to overcrowding on trails. The ATC found that significantly more thru-hikers on the AT register as northbound hikers with a southern terminus start. Far less register as southbound hikers with a northern terminus start. Even fewer register as flip floppers, where hikers start in the middle of the trail, hike toward one terminus and go back to where they started to hike toward the other terminus to finish. These options for starting an AT thru-hike can be found in Figure 4.6 [15]. However, the exceptionally high demand for campsites along the Appalachian Trail can threaten its resources. The flip flop itinerary for thru-hiking (start midway through the trail and hike in one direction, transport back to midway start point and hike in the opposite direction [15]) evenly distributes the use of Appalachian Trail campsites and reduces the aggregate camping impact and crowding, thus alleviating stress on trail resources. For this reason, the ATC recommends flip flopping for thru-hiking the AT. “The increase in flip-flop thru-hikes indicates that many hikers have recognized the benefits of non-traditional thru-hiking itineraries, both for themselves and the Trail ”[16].



Figure 4.6: Options for starting an Appalachian Trail thru-hike [15]

The Appalachian Trail Conservancy also strives to protect the “trail experience”. The Appalachian Trail Experience is defined to represent “the sum of opportunities that are available for those walking the Appalachian Trail to **interact with the wild, scenic, pastoral, cultural, and natural elements** of the environment of the Appalachian Trail, unfettered and unimpeded by competing sights or sounds and in as direct and intimate a manner as possible.” In the July/August 1997 edition of the *Appalachian Trailway News*, the summary definition included the following opportunities [41]:

- Opportunities for **observation, contemplation, enjoyment, and exploration** of the natural world.
- A sense of remoteness and detachment from civilization.

- Opportunities to experience solitude, freedom, **personal accomplishment, self-reliance, and self-discovery.**
- A sense of being on the height of the land.
- Opportunities to experience the **cultural, historical, and pastoral elements** of surrounding countryside.
- A feeling of being **part of the natural environment.**
- Opportunities for travel on foot, including opportunities for **long-distance hiking.**

In studying the Appalachian Trail Experience, we noted the phrase, “interact with the wild, scenic pastoral, cultural, and natural elements,” and we associated that with the Appalachian Trail's mission statement in Table 4.1. We also note key words such as “observation”, “contemplation”, “self-discovery”, “cultural”, “historical”, and “natural environment”. Combined with the Appalachian Trail's mission statement and Karen Berger's characterization, we gather that the Appalachian Trail has a culture of hikers reflecting on themselves and experiencing self-realization from the heritage and natural beauty of the trail. The practice of Leave No Trace and other conservation-conscious practices is imperative to maintain the this trail experience and sustain this culture.

The Pacific Crest Trail Association strives to not only practice conservation to protect the trail, but also to protect the “trail experience”. “Trail users seek the tangible and intangible benefits of wandering amongst the **exceptionally scenic, wild, natural, and historic landscapes** along the crest of the Pacific ranges of the United States. The Trail experience, as used in this context, represents the sum of experiences available to the traditional and intended user traveling along the PCT” [1]. The PCT experience is described to focus

on “natural” and “undisturbed” landscapes, so the need to practice Leave No Trace is imperative. We noted that the PCT's mission statement also demonstrates the importance of maintaining a “world-class experience” and the value of maintaining “wild and scenic lands.” In recalling Karen Berger’s description of the Triple Crown trails, we realize that an important piece of the Pacific Crest Trail's culture is embracing the wilderness of the PCT and marveling at the drastic changes of terrain along the trail.

Chapter 5

Experiments

5.1 Experiment to Describe the Triple Crown Trails

In our first round of experiments, we needed to find meaningful topics in the Triple Crown communities. We used LDA to identify topics discussed in the AT community, PCT community, and CDT community. In this experiment, each trail community's corpus of tweets was composed of tweets relating to its respective trail. The tweets from these experiments were collected from January 2017 to May 2017. The seeds used for tweet collection can be found in Figure 5.1. Using `LDAWrapper`, we cleaned our tweets, removed stopwords, URLs, punctuation, and the RT marker which indicates that the tweet was retweeted. We then made all of the tweets lowercase so that terms like “patagonia” and “Patagonia” would be treated as the same term. Then we ran LDA over 100 iterations to look for the appropriate number of topics and word distributions. Running fewer iterations resulted in sloppier word distributions, and running more than 100 iterations resulted in `OutOfMemory` errors. This number of iterations was recommended by the Spring CS 5604 Topic Analysis team, and it produced intelligible topic results [37]. We then found that 7 topics of 7-word distributions were best for our dataset, because 5 topics on 5-word distributions were too condensed and unrepresentative of the tweets in our dataset while 10 topics of 10-word distributions held a lot of repetition in terms. When testing 7 topics of 7-word distributions, we found topics that were meaningful to us. The results of this experiment can be found in Figure 5.2.

Using rawgraphs.io, we were able to create visualizations for the Appalachian Trail topics and Pacific Crest Trail topics [35]. We created treemap visualizations, where each grouping of blocks of the same color are a topic, and the size of the block is correlated to its term's probability of occurring in the topic. We also created circle packing visualizations, which also uses size to represent every term's probability of occurring in the topic relative to each other. Each cluster of shapes represents a topic, but in the circle packing visualization, it is easier to visually compare the probabilities that each topic occurs in the dataset. A cluster of circles that occupies more space means that that topic is likely to occur in more tweets than that of a cluster that occupies less space in the visualization.

Trail	Number of Tweets	Search Queries
Appalachian Trail	24,954	“Appalachian Trail” “#appalachiantrail” “#at2017” @AT_Conservancy
Continental Divide Trail	2,508	“CDT” “#cdt” “#cdt2017” @CDNST1 “#continentaldivide trail” “Continental Divide Trail”
Pacific Crest Trail	9,444	“Pacific Crest Trail” “#pacificcresttrail” “PCT” “#pct” “#pct2017” @PCTAssociation”

Table 5.1: Details of tweet collection

From Figure 5.2, we can see that the Appalachian Trail community leans heavily toward discussing their in-the-moment experiences on the trail. AT Topic 1 refers to McAfee Knob—the most photographed site on the Appalachian Trail [42]. AT Topic 2 is suggestive of watching

the sunrise over the Catawba Valley on a backpacking trip. AT Topic 5 describes a woman's hiking trip.

Conversely, the Pacific Crest Trail tweets related more to planning the logistics of a hike. For example, PCT Topic 1 was about resupply tips in California, and PCT Topic 4 was about acquiring Danner boots at an Outdoor Research show. PCT Topic 6 is about hikers taking months off from their jobs for their PCT backpacking trips.

The Continental Divide Trail's datasets seemed to yield topics relating bravery on the CDT and supporting the Continental Divide Trail Coalition (Twitter account: @cdnst1). However, the terms of the CDT's topics were repeated through the seven topics. When looking at the datasets, the Appalachian Trail's corpus had 24,915 tweets and the Pacific Crest Trail's corpus had 8,547 tweets. On the other hand, the Continental Divide Trail's corpus had 2,457 tweets—less than 10% of the Appalachian Trail's and less than 35% of the Pacific Crest Trail's.

Topic #	AT Topics	PCT Topics	CDT Topics
Topic 1	#indigenous, #tairp, #amerianindian8, day, knob, mcafee, trailva	California, #pct2017, 2, story, tips, resupply, #pics	help, #bravethecdt, #hikecdt, today, @cdnst1, vote, great
Topic 2	va, catawba, sunrise, halfway, #backpacking, just, oc4444x2400	@pctassociation, like, today, #lwcf, win, did, great	#bravethecdt, @cdnst1, #hikecdt, help, great, support, today
Topic 3	days, amp, long, mountain, complete, miles, week	mount, adams, goat, rocks, @hogansog, washington, view	#bravethecdt, help, today, @cdnst1, support, vote, great
Topic 4	#travel, #bestseller, black, 1, awol books, 2	#orshow, booth, gear, come, free, #pct2017, @danner	#bravethecdt, help, today, great, @cdnst1, support, #hikecdt
Topic 5	new, going, woman, 80yearold, solo, sisters, twin	wild, lost, #travel #bestseller, oprahs, #7, #8	#bravethecdt, #hikecdt, @cdnst, support, help, great, today
Topic 6	hiker, #at2017, @thetrekat, 5, update, thruhiker, thruhikers	taking, #backpacking, months, job, better, 4, day	#hikecdt, #bravethecdt, @cdnst1, great, help, support, today
Topic 7	#at2017, @thetrekat #trail, gear, list, things, #photography	#pctdays, new, instagram, year, weeks, posts, bitesized	help, #bravethecdt, today, #votecdt, vote, 25k, @cdnst1

Table 5.2: Results of topic analysis of tweets: January-May

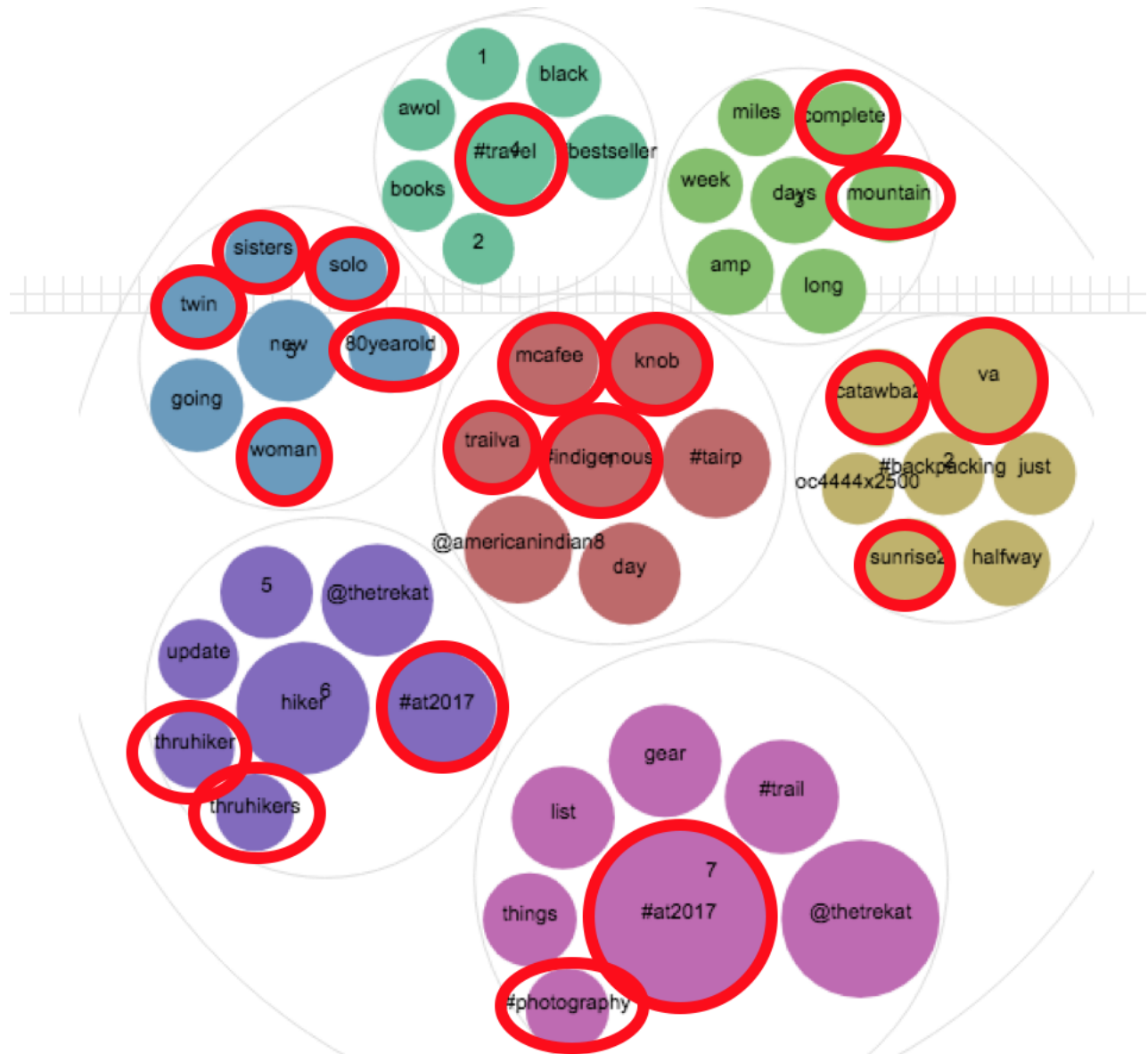


Figure 5.1: Circle packing visualization of topic analysis on Appalachian Trail tweets

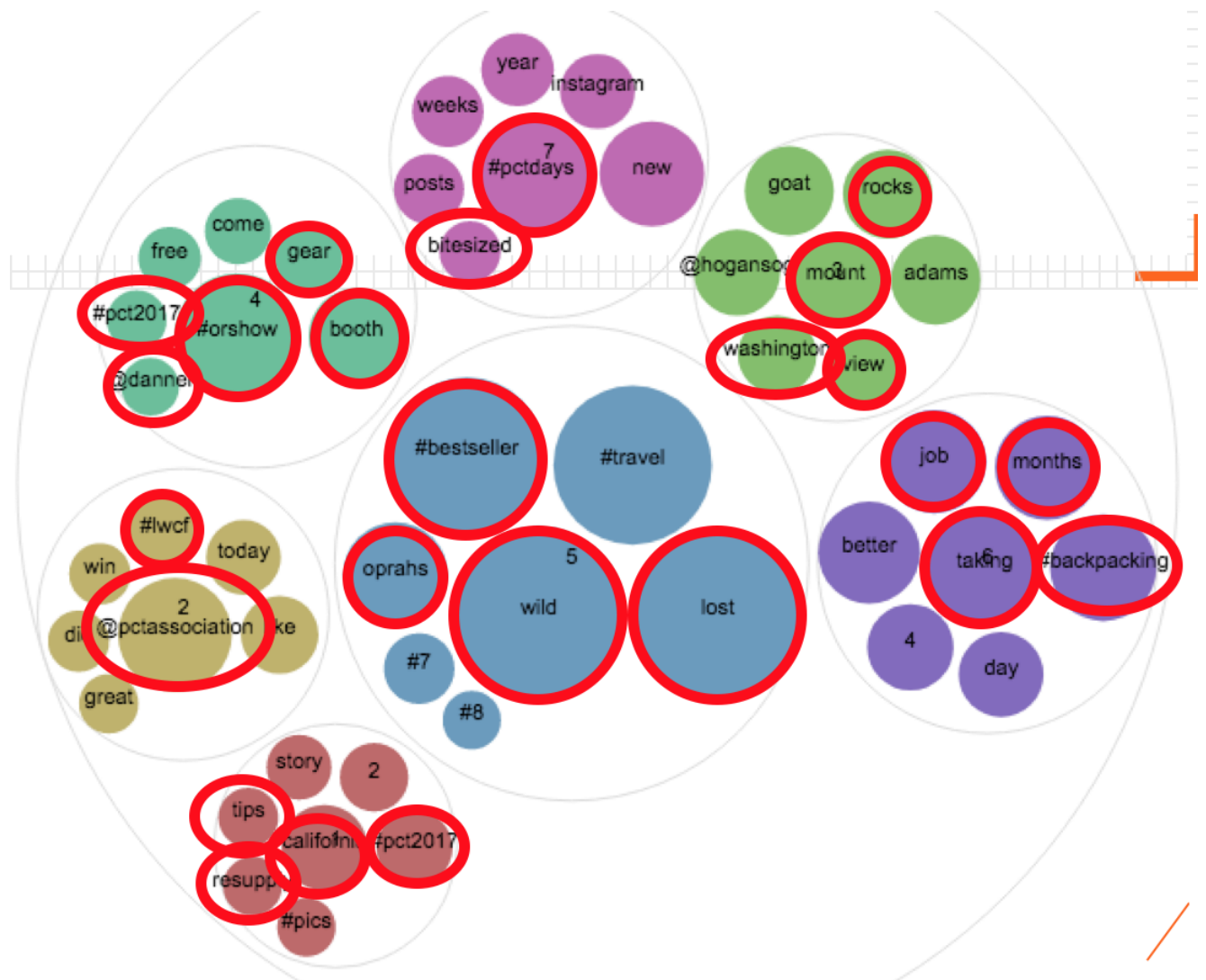


Figure 5.2: Circle packing visualization of topic analysis on Pacific Crest Trail tweets

5.2 Experiment to Explore Depreciative Behavior

The United States National Park Service defines depreciative behaviors as “actions that degrade park resources or the experiences of others”. Specific examples of depreciative behavior are [43]:

- littering



Figure 5.3: Treemap visualization of topic analysis on Appalachian Trail tweets

- feeding of wildlife
- collection of green wood for use in campfires
- disfiguring trees and rocks
- improper disposal of human waste in backcountry
- walking and hiking with pets that are not leashed
- collection of specimens

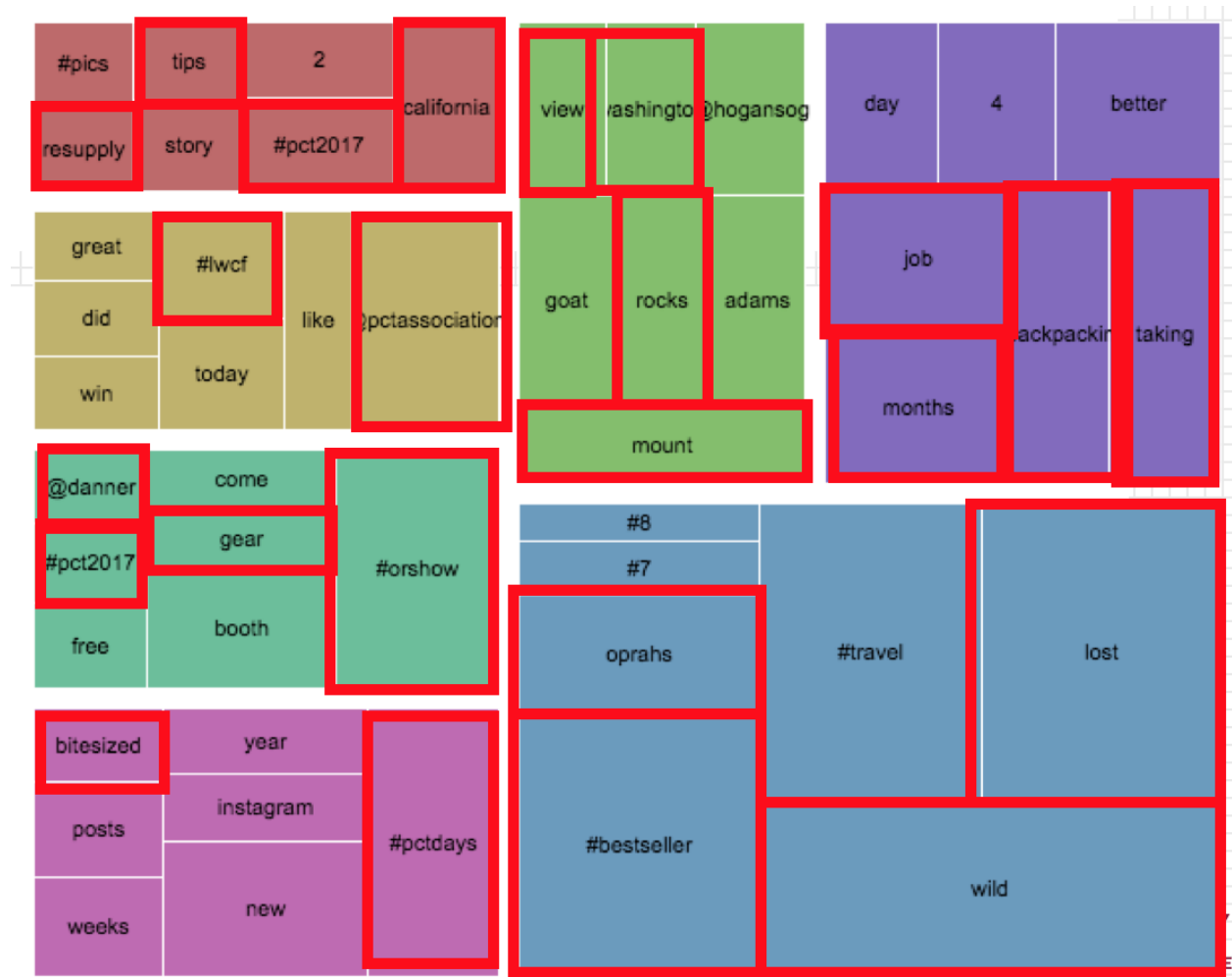


Figure 5.4: TreeMap visualization of topic analysis on Pacific Crest Trail tweets

- construction of fire rings and ground fires in the backcountry
- shortcutting on trail switchbacks

Littering seemed to be an obvious form of depreciative behavior as children are taught from a young age to not litter. However, collecting specimens, like rocks and flowers, and shortcutting on trail switchbacks, are less obvious forms of depreciative behavior since their environmental impacts are less evident to those who do not spend much time on hiking trails. We decided that an interesting study might be to see which forms of depreciative behavior

were most discussed on Twitter. Using a dataset of 1,592,059 tweets we built a classifier to identify which tweets described a specific depreciative behavior. Table 5.3 provides details of this corpus.

Seed	Description	Number of Tweets
AnishHikes	tweets pulled from a Triple Crown veteran's Twitter Account	3,082
Appalachian Trail	search query for “Appalachian trail”	71,324
AT2017	search query for “at2017”	711
@AT_Conservancy	tweets pulled from Appalachian Trail Conservancy 's Twitter account	3,797
@CDNST1	tweets pulled from Continental Divide Trail Coalition's Twitter account	2,291
CDT2017	search query for “cdt2017”	17
Continental Divide Trail Hike	search query for “continental divide trail” search query for “hike”	1,779 1,477,708
@hikewithgravity	tweets pulled from AT hiker who plans to hike PCT	543
@leavenotrace	tweets pulled from the Leave No Trace Twitter account	3,266
Pacific Crest Trail	search query for “pacific crest trail”	21,639
PCT2017	search query for “pct2017”	378
@PCTAssociation	tweets pulled from Pacific Crest Trail Association's Twitter account	878
@PCTNews	news posted by Halfmile PCT maps	1,261
@O_Cliff23	tweets pulled from a Leave No Trace trainer who has hiked AT	213
@walkingthecd	tweets pulled from hikers who have hiked the CDT	172

Table 5.3: Details of tweet collection—one year of collecting

Labeling training data is integral to building a classifier. In hand labeling the training data for this classifier, we were able to intimately familiarize ourselves with tweets that mentioned depreciative behaviors. In the fall semester of 2017, Chon et al. built a tweet visualization tool for the CS 5604 class that made hand labeling the training data significantly easier, as it enabled us to search for key words relating to the depreciative behavior (e.g., searching for trash would correspond to littering) [11]. As shown in Table 5.4, littering and walking with pets off-leash are the most tweeted about depreciative behaviors.

Depreciative Behavior	Number of Tweets
Littering	55
Feeding of wildlife	8
Collection of green wood for use in campfires	0
Disfiguring trees and rocks	28
Improper disposal of human waste in backcountry	36
Walking and hiking with pets that are not leashed	54
Collection of specimens	10
Construction of fire rings and ground fires in the backcountry	4
Shortcutting on trail switchbacks	1

Table 5.4: Training data—breakdown of tweets for depreciative behaviors

Depreciative Behavior	Number of Tweets
Littering	55
Walking and hiking with pets that are not leashed	54
Improper disposal of human waste in backcountry	36
Disfiguring trees and rocks	28

Table 5.5: Depreciative behavior categories with more than 10 tweets

We built a single multi-class SVM classifier using the training data from Figure 5.4. This classifier was built in two phases—labeling the training data and training the SVM classifier. The data was labeled by the author of this paper, with assistance from the Tweet Visualization team in the CS 4784 Technology on the Trail Capstone course [21][36]. This division of labor was to avoid any question of developer bias in building the classifier. During the hand labeling efforts, we were able to intimately study the tweets in the dataset. We wanted to capture every motivation to tweet about trail depreciation—examples of good behavior that counteract trail depreciation (e.g., Figure 5.5), advocacies against depreciative behavior (e.g., Figure 5.6), and showcases of depreciative behavior (e.g., Figure 5.7).

Since our corpus had few tweets relating to feeding wildlife, collecting green wood for campfires, collecting specimens, constructing fire rings and ground fires in the backcountry, and shortcutting on trail switchbacks, our classifier excluded these; we did not have enough training data. This was justified based on an understanding that the aforementioned depreciative behaviors are less known to those that do not study wilderness conservation. Feeding wildlife, at first glance, does not seem to be so awful since some people enjoy feeding ducks in city parks. However, after considering the consequences of feeding bears and wild boars,

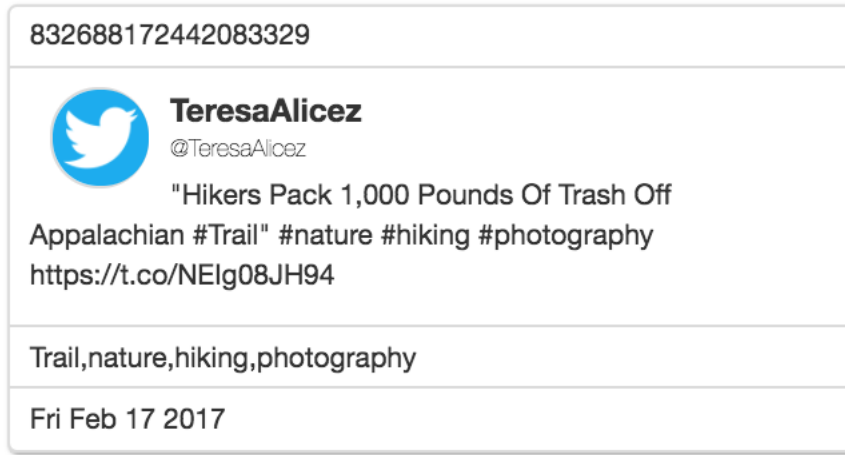


Figure 5.5: Example of good behavior on a hike to counteract trail depreciation

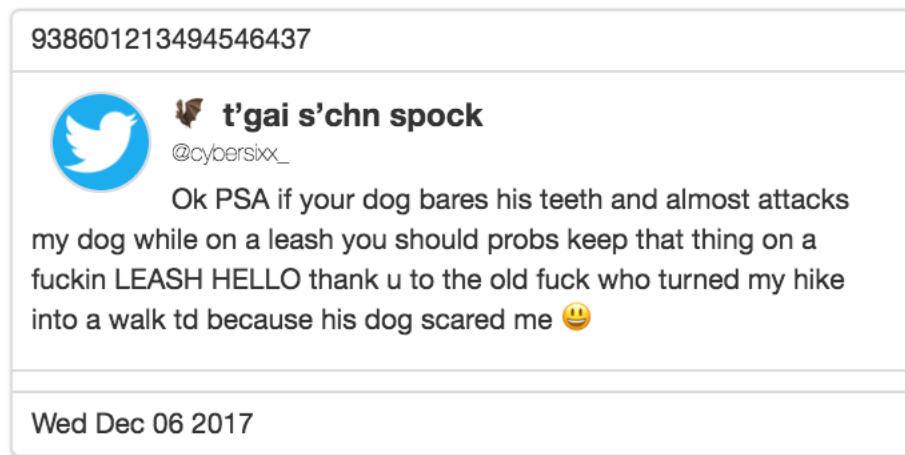


Figure 5.6: Example advocating against trail depreciative behaviors

it becomes evident that encouraging them to grow comfortable with human contact can actually be dangerous to hikers. Similarly, the consequences of collecting green wood for campfires, collecting specimens, constructing fire rings and ground fires in the backcountry, and shortcutting on trail switchbacks, are not obvious in their depreciative impact on trails. So, our classifier focused on categorizing tweets for describing littering, disfiguring trees and rocks, improper disposal of human waste, and walking pets off-leash.



Figure 5.7: Example of showcasing trail depreciative behavior

When run against our “Hike” dataset, our classifier performed well in identifying tweets about littering and walking pets off-leash. Our evaluation was done manually, and the performance results of the classifier can be found in Table 5.6. However, the classifier treated labels for disfiguring trees and rocks and improper disposal of human waste as a “catch-all” for tweets that may not necessarily have been related to the respective labels, but were certainly not related to littering or walking pets off-leash. Since there were thousands of tweets that were incorrectly classified as improper disposal of human waste, we chose a random sample of 194 tweets to evaluate. We came to the conclusion that hikers on Twitter are likely less concerned with disfiguring elements in nature or the proper disposal of human waste. This is a probable conclusion since the consequences of littering and the presence of unleashed pets are more prevalent on trails, in comparison to how often graffiti or the candid disposal of human waste are witnessed on a typical hike.

Depreciative Behavior	Number of Tweets	Tweets Correctly Classified
	with over 90% Probability	
Littering	67	60/67 (90%)
Walking and hiking with pets that are not leashed	35	29/35 (83%)
Disfiguring trees and rocks	180	30/180 (17%)
Improper disposal of human waste in backcountry	6,335	27/194 (14%)

Table 5.6: Manually evaluating the performance of the classifier

5.3 Validating Depreciative Behaviors by Exploring Blog Posts

Dr. Edward Fox taught CS 4624 (Multimedia, Hypertext, and Information Access), a capstone class where each team was able to spend the semester serving a client on a project relating to multimedia information collection and hypertext processing. As a client, the author of this thesis was able to guide a team to gather data that could help in this thesis. The Trail Study team was very useful in collecting blog posts from hiking blogs. This data was useful because it provided an opportunity to compare the trends found in tweets from the aforementioned experiment in Section 5.2 to blog posts of the trail community. The team fed 31 hiking blogs into a web crawler to store text, data, and URLs into a database. They implemented an automatic tagging system that generates tags based on the content of the blogs and integrated that tagging system into an interface that allowed us to search through

the blog posts [20]. We were able to search the posts for terms that related to a depreciative behavior. For example, to find blog posts about littering, we searched for posts that had the words “litter”, “garbage”, “trash”, “waste”, “throw”, and/or “toss”. To find blog posts about disfiguring trees and rocks, we searched for posts that had the words “trees”, “rocks”, “drawing”, “draw”, “drew”, “paint”, “painting”, “painted”, “art”, “ruin”, and/or “graffiti”. We then read through the blog posts that were returned for each behavior and determined whether or not that blog post related to its respective depreciative behavior. It should also be noted that some blog posts, especially those by LNT advocates, could count for multiple depreciative behaviors. This resulted in a dataset with 3,459 unique blog posts.

During the experiment to explore depreciative behavior, we found that the most frequently tweeted depreciative behaviors were littering and hiking with pets off-leash. However, this was not the case for the collection of blogs. Table 5.7 lists the number of blog posts that reflected each depreciative behavior. While the number of blog posts for each depreciative behavior is not nearly as high as the number of tweets for littering and hiking with pets off-leash, the numbers of blog posts representing each depreciative behavior are more evenly distributed. That is to say, with the exception of littering, collecting green wood for campfires (0 tweets and 0 blog posts), and shortcutting trail switch backs (1 tweet and 0 blog posts), the depreciated behaviors were represented in 0.05%–0.24% of blog posts. Meanwhile, at best 0.0004% of tweets represented littering, and all of the other depreciative behaviors were represented less. We reflect on why the representations could be different between tweets and blog posts in Chapter 6 of this thesis.

Depreciative Behavior	Number of Blog Posts
Littering	35
Feeding of wildlife	2
Collection of green wood for use in campfires	0
Disfiguring trees and rocks	6
Improper disposal of human waste in backcountry	5
Walking and hiking with pets that are not leashed	5
Collection of specimens	4
Construction of fire rings and ground fires in the backcountry	8
Shortcutting on trail switchbacks	0

Table 5.7: Blogs mentioning depreciated behavior

5.4 Visualizing Tweets on the Trail

In the Spring 2018 semester, Dr. McCrickard taught CS 4784, Technology on the Trail, a senior capstone that “examines and defines ways that technology can influence hiking and outdoor experiences—both in positive and negative ways—for individuals and groups” [36]. This class was split into teams, where each team was assigned a term project. The author of this thesis served as a client for the Tweet Visualization team. The Tweet Visualization team's initial step was to create a word cloud visualization to see which terms frequently occurred in the dataset described in Table 5.3. After cleaning the data of duplicated tweets and removing tweets that used the word “hike” in a political sense (i.e., tax hike), they were able to produce the word cloud found in Figure 5.8. The word cloud was useful because it

5.12 than there are in Figure 5.11.

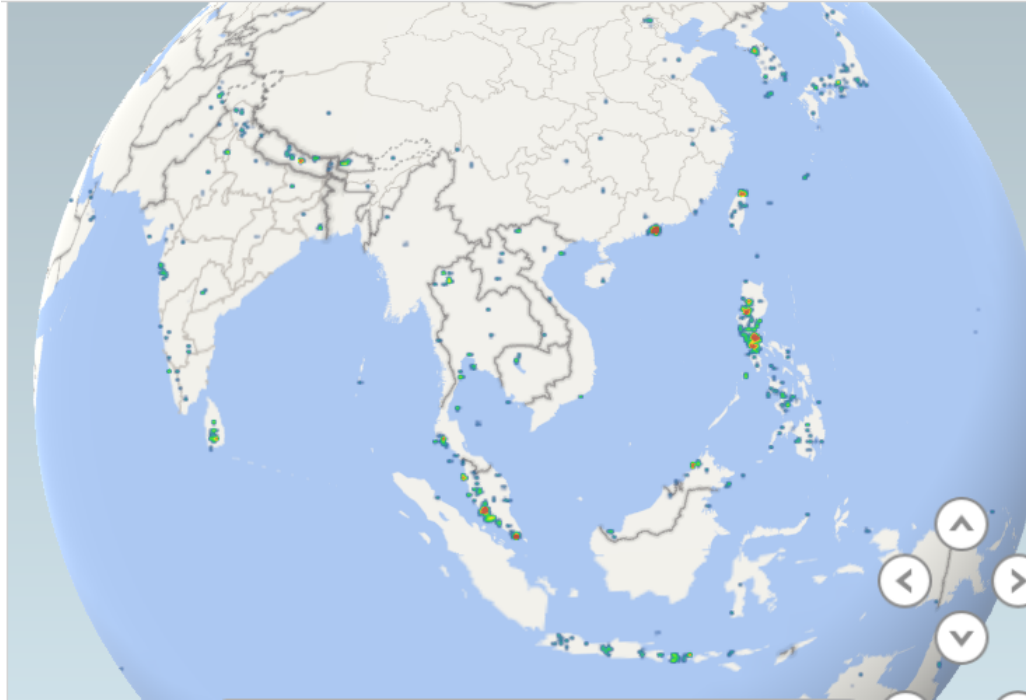


Figure 5.9: Heatmap of tweets in Asia [21]

From the visualization in Figure 5.11, we saw that the majority of the tweets were from the Appalachian Trail and the Pacific Crest Trail, which was expected. The Appalachian Trail “is the outdoorsperson’s Holy Grail” and “is the trail that people choose for their trip of a lifetime” [7]. The Pacific Crest Trail is the second national scenic trail in the United States and holds the moniker of “second twin” and “younger sibling” to the Appalachian Trail [7]. Given the prestige of these trails, it makes sense that there would be a higher density of tweets near those trails. We predicted that spring, summer, and fall would have the most tweets since a hiking season is characterized as spring–fall. Needless to say, we were surprised to see so many hikes from the winter season. That is not to say that the number of winter tweets dominated the number of tweets from the other three seasons, but we did anticipate that the spring and summer seasons would have more tweets. We will discuss this in Section

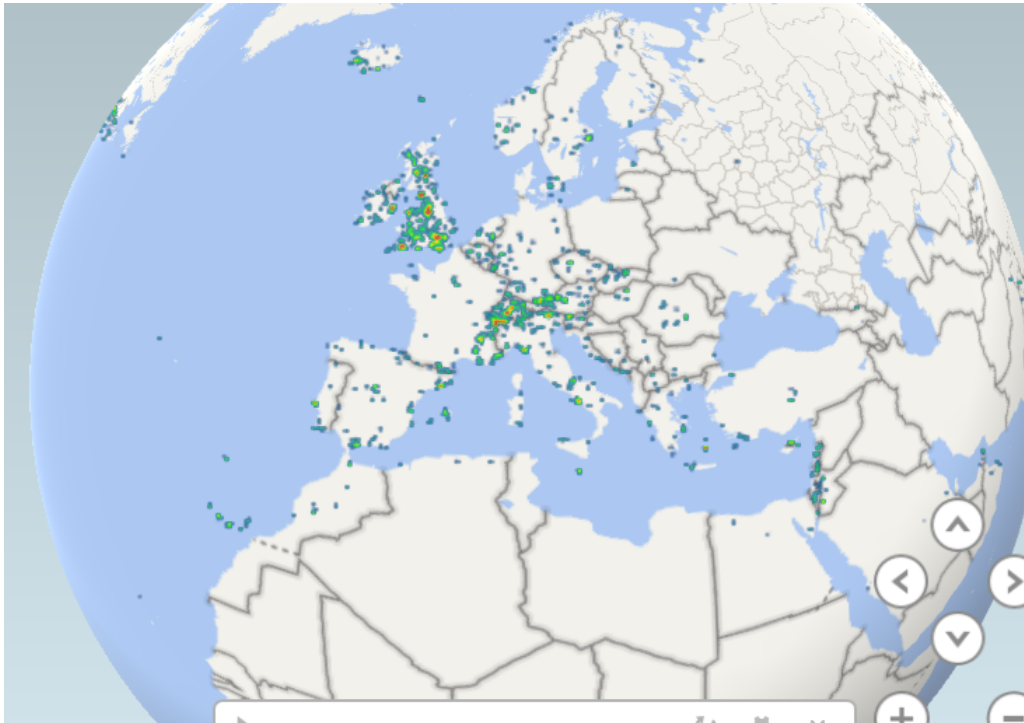


Figure 5.10: Heatmap of tweets in Europe [21]

6.3 of this thesis.

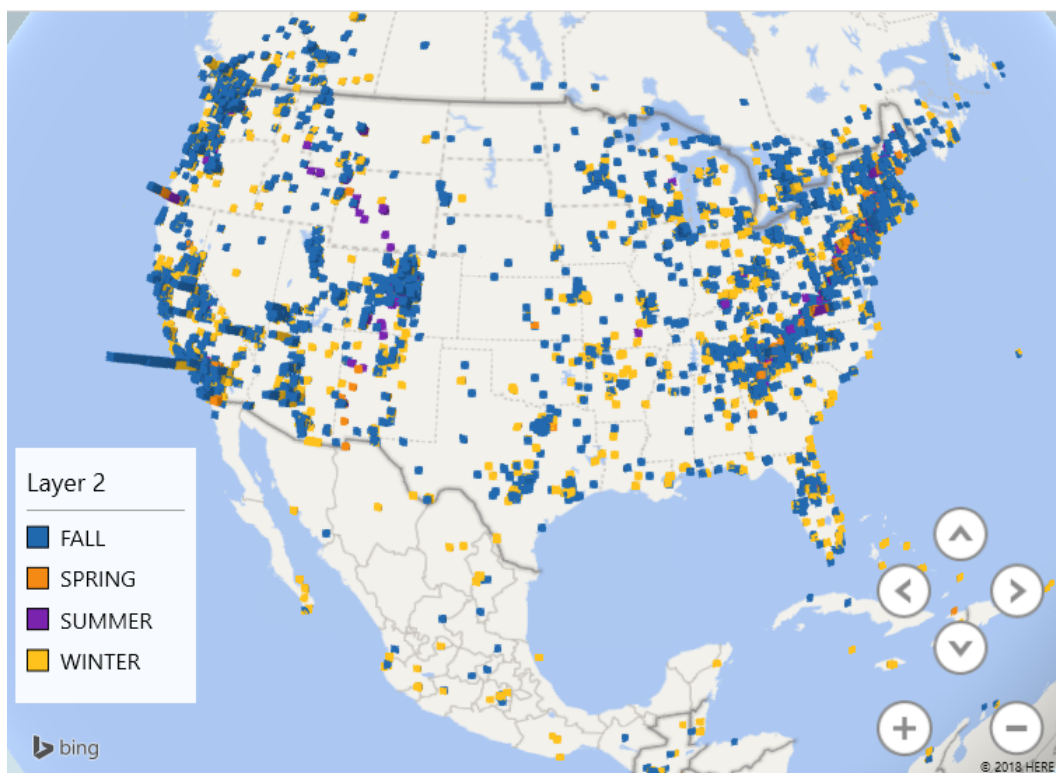


Figure 5.11: Map of tweets from the United States with a layer turned on to mark the season during which it was tweeted [21]

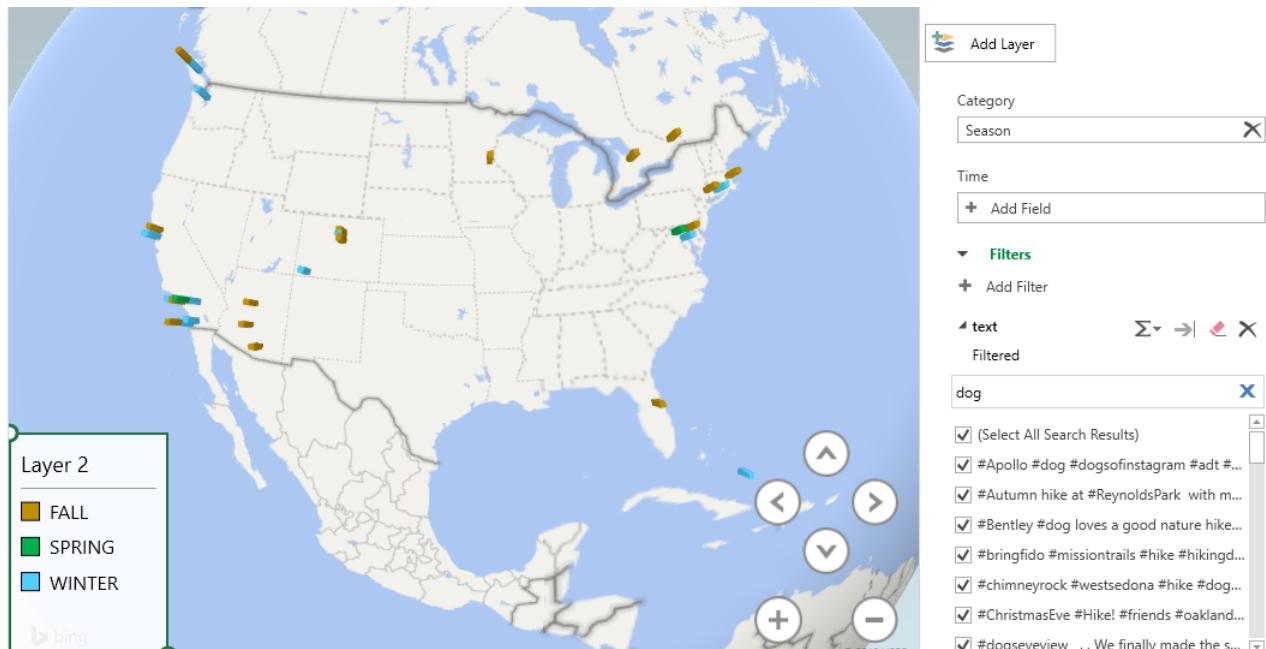


Figure 5.12: Map of tweets from the United States with a filter that only displays tweets containing the search term “dog” [21]

Chapter 6

Conclusion

The notion of trail culture is one that is simultaneously predictable and nebulous. On the one hand, hiking is arguably one of the most accessible outdoor recreation activities. In its purest definition, a hike is a vigorous walk, so prior experience and equipment are not necessarily prerequisites. However, there are also hikers who go through intense physical training and invest much money into quality hiking gear in preparation for a major hike [32] (Table 4.2). With hikers approaching the trail for hiking activities that are so diverse in nature, it is difficult to give a generalized description of trail culture. The purpose of this thesis was to explore the different cultures within the larger umbrella of “trail culture” and give those cultures meaningful descriptions.

6.1 Results

6.1.1 RQ1: What are the cultural differences among the Triple Crown Trails?

From Sections 4.1 and 5.1, we were able to find distinct characteristics that described the Appalachian Trail's and Pacific Crest Trail's respective cultures. The Appalachian Trail community is a “community in the wilderness”. It encourages each hiker to find inspiration in its natural beauty and heritage so that they may have meaningful experiences on the

trail. The Pacific Crest Trail is the “soul of American wilderness”. Its community values its world-class hiking experience, and the tweets reflect that by showing topics that demonstrate the meticulous planning of a hike.

What can tweet analysis do to describe trail cultures?

We can analyze large datasets of tweets by using natural language processing to validate observations found through qualitative analysis.

What causes a hiker to choose a particular trail?

Location plays a large factor in deciding which trail to hike. We spoke with an Appalachian Trail thru-hike veteran who told us that she chose the Appalachian Trail because she grew up on the east coast and near the Appalachian Trail. This sentiment is shared by many thru-hikers and section hikers alike. However, some hikers actually choose their hikes based on the hiking experience. Since the Appalachian Trail is the oldest of the Triple Crown trails, it is the most well-known, and that draws the attention of a lot of hikers when it comes to choosing a trail. Many hikers, of all levels, will attempt the Appalachian Trail because of its prestige. Others are drawn to the strong sense of community on the trail—the feeling that they will not be walking alone for long. Similarly, the Pacific Crest Trail draws a crowd because of its changing terrain. The trail starts in the desert, on the Mexican border, and as it goes farther north, it turns into the High Sierras, and then the dense forest until it reaches Canada. Some hikers crave the changing terrains and challenges that come with surviving in the desert and making their way through the snowy mountains of the High Sierras. Some hikers want the solitude of hiking alone, and they only seek socialization on occasion.

6.1.2 RQ2: What specific depreciative behaviors do hikers tweet about more than others?

In Section 5.2, we found that litter and graffiti are discussed the most in tweets. In Section 5.3 we found litter was discussed in blog posts far more often than any other depreciative behavior.

What categorizations of tweets represent depreciative behaviors on the trail?

The tweets about depreciative behaviors can be categorized into three groups: tweets that advocate against depreciative behaviors, tweets that are evidence of engaging in or supporting depreciative behaviors, and tweets that show evidence of depreciative behaviors.

Are tweets about trail depreciation more focused on advocating against, engaging in, or showing evidence of trail depreciation?

This depends entirely on the depreciative behavior in question. The tweets about littering were mostly advocating against littering or showing evidence of other visitors who had littered. The tweets about walking with your pets off-leash were split between advocating against walking with dogs off-leash and showing support or engagement of dogs walking off-leash.

6.2 Discussion

One contribution of this thesis was that it identified the cultural differences between the Appalachian Trail, Continental Divide Trail, and Pacific Crest Trail. The AT is a “community in the wilderness,” whose conservancy organization values “natural beauty and priceless cul-

tural heritage”. Its tweets reflect this value of beauty with topics that capture in-the-moment experiences. The PCT is the “soul of American wilderness” whose conservancy organization is devoted to providing its hikers with “a world class experience”. By running topic analysis on these tweets, we saw that the PCT valued meticulously planning the logistics of a hike in order to achieve this “world class experience.” The CDT, whose sobriquet is “the wild child,” did not have enough CDT tweets to effectively run topic analysis. This was not entirely surprising since the CDT is the newest of trails, and it still has yet to have a clear, marked hiking trail [7] [6].

Another contribution of this thesis was that we were able to identify which forms of trail depreciation were discussed the most. We found that on Twitter, littering and hiking with animals off-leash were tweeted about significantly more than the other forms of trail depreciation. However, when we looked for blog posts that mentioned depreciative behaviors, we found that littering was the most discussed depreciative behavior, by a landslide. After that, the other behaviors were discussed rather evenly. When it came time to see if the tweets and blog posts were written in advocacy against depreciative behaviors, or if they were written as evidence of engaging in depreciative behavior, we found that the medium of communication was significant. Figure 5.5 shows an example of a tweet that demonstrates good behavior on the trail to counteract a depreciative behavior, while Figure 5.6 advocates against a depreciative behavior. However, there are also tweets like Figure 5.7 that showcase the author's participation in depreciative behavior. In contrast, the blog posts that mention depreciative behaviors are all in advocacy against depreciative behavior, and many encourage Leave No Trace practices.

Zhao and Rosson stated that people tweet because of its informal nature, which allows people

to provide brief and easy updates to others, the ability to share information and ideas as they occur in real time, the brevity of the character limitation, and the lack of required effort to read tweets and post tweets that comes along with the character limitation [46]. On the other hand, Nardi et al. likened blogs to diaries and personal journals. They said that bloggers are motivated because they want to: “update others on activities and whereabouts, express opinions to influence others, seek others' opinions and feedback, “think by writing”, and/or to release emotional tension” [39]. Blogging requires more thought and effort to fulfill these motivations because it is not as simple as a 280 (or 140, before 2017) character thought. It is more committal because it is implied that the author is writing with intent. To blog about something as serious as trail depreciation, there has to be a strong motivation behind the blog post. The passion to care for the outdoors is strong enough to motivate a blog post. It is less probable that someone cares deeply enough about harming the trail and the environment around it to actually blog about it. It is more likely that someone partaking in depreciative behavior would not care enough about the behavior to write a blog post on it. So, we see blogs can represent the hiking community differently than tweets.

6.3 Limitations

This study is not without its limitations. Ultimately, this thesis studies the cultures and values of the hiking community, and much of our data comes from the community's presence on Twitter. A community is most valuable when all of the members are working to maintain the community [17], but this strength is not present in the presented community. Part of this is because at the end of the day, hiking is a recreational activity in which anyone can partake, and to partake in the activity is to be part of the community. However, you cannot mandate that everyone in this already large community actively contribute to the discus-

sion by defining the values and motivations of such a broad community with such broad motivations. Similarly, not everyone in the hiking community is on Twitter, and there are even fewer members of the hiking community who are openly tweeting about their hiking motivations and values.

Similarly, most tweets do not contain their geographic coordinates because many Twitter users opt out of sharing their location. This means that the PowerMap visualizations are not a complete representation of the dataset.

Another limitation of this study was in the collection of tweets. The seeds for collecting tweets did not all start harvesting at the same time. In January, we only harvested the seeds relating to the Triple Crown trails and their conservancy organizations. Collection using the seeds for specific hikers's Twitter accounts and the general search query for “hike” did not begin until October. Also, there were many tweets that were harvested from the search query for “hike” that were actually about tax hikes, but we did make an effort to filter out those tweets. However, they do make up part of the 1.5 million count.

Also important to note is that the author's reflections about hiking and hiking preparation were collected from her own experiences as a recreational, albeit avid, hiker. There was no formal ethnographic study of interviewing stakeholders; however, it is important to note that the preliminary studies of conservation culture and the research of depreciative behavior were executed with guidance from Dr. Jeffrey Marion—a recreation ecologist, founder of Leave No Trace, adjunct professor at Virginia Tech's Department of Forest Resources, and committee member for this thesis.

6.4 Future Work

This thesis was an exploratory study that showcases the ways that we may use natural language processing techniques to understand and describe the cultures on hiking trails. We hope that this thesis can lead to more natural language processing studies for technology on the trail and human-computer interaction outdoors research. For example, this thesis specifically focuses on hikers, but there are many other stakeholders of trails. There are educational groups, like the Boy Scouts of America, there are park rangers and ridge runners, and there are emergency rescue teams. The studies we performed for this thesis could similarly be applied to any of these stakeholders, instead of hikers. Conversely, more classifications could be made to predict the different types of hikers (i.e., day hikers vs. section hikers vs. thru-hikers). We also recommend a more in-depth approach to qualitative research (i.e., surveys, focus groups, interviews, ethnographic studies in hiking groups, and case studies).

Those interested in continuing this research may want to place greater emphasis on studying the statements on trail experience over the mission statements. During the defense for this thesis, Dr. Jeffrey Marion gave caution to read less into the mission statements of the trail conservancy organizations. He posited that hikers may not necessarily read the mission statements of a trail when deciding to embark on a backpacking trip. Furthermore, as a member of the Appalachian Trail Stewardship Council, he stated that the council uses the official statements on trail experience to make its decisions. He said that they are less concerned with the mission statement. It would be interesting to replicate the experiments for RQ1, where the documents describing trail experiences replace the mission statements as the focus for the qualitative study. Results from such a study would demonstrate how the mission statements differ from the trail experience documents.

Another avenue for furthering this research effort would be to conduct a similar study, where the quantitative analysis is performed on blog posts and trail journals instead of tweets. Due to the brevity of tweets, blog posts and trail journals are more descriptive and allow for a deeper exploration of the writers' thoughts. Since blog posts and trail journals would inherently be more detailed, we could gain more insights to what the thru-hikers value about being on the trail. While tweets are intended to convey real-time, in-the-moment ideas, blog posts and trail journals convey ideas that have been further developed and were written with intention. Performing topic analysis and classification with blog posts and trail journals could provide different results and highlight the differences between hikers who tweet and hikers who blog or journal.

Dr. Jeffrey Marion's research seeks to understand a “small group of thru-hikers, usually younger, who are immune to the Leave No Trace concept and education”[31]. That is to say that this group engages in depreciative behaviors. According to Dr. Jeffrey Marion, this group hikes with the mantra “hike your own hike,” and they use that mantra as an excuse to not abide by the principles of Leave No Trace. This group is very difficult to identify via tweets for two reasons. First and foremost, this group makes up a small percentage of hikers and the tweets in our dataset did not reflect this group's presence. Secondly, the phrase “hike your own hike” has many meanings. Joe Hikes, who completed his PCT thru-hike in 2014, reflects on the various meanings of “hike your own hike”. He provides the following interpretations of the mantra [26]:

- Do things your own way; don't try to copy someone else.
- Don't try to tell other hikers how to hike.
- Feel free to set goals for yourself that feel like a good fit, even if they're unusual.

- Don't always change for the sake of the group; maintain your individual interests first.

It appears that the interpretations that best reflects the group that Dr. Jeffrey Marion talks about is “don't try to tell other hikers how to hike” and “don't always change for the sake of the group; maintain your individual interests first”. This group does not want to be told how to hike, and they refuse to change just because they are told to practice Leave No Trace. However, there are also interpretations that just mean to hike at your own pace even if others have hiked at quicker paces, or to hike with a trail weight that works for you regardless of how much or little others are carrying. There are interpretations that encourage hikers to feel proud of the goals that they have accomplished, even if they differ from the goals of other hikers. Joe Hikes writes the following. “There is a lot of hype out there about how shameful it is to “fail” your thru-hike attempt and how you should avoid it at all costs. I think it's nonsense. You should go wherever you feel like you need to go, and as long as you know that you're still pursuing new opportunities for personal growth, you don't have to feel any shame about leaving the trail” [26]. These interpretations of “hike your own hike” do not necessarily represent hikers who refuse to practice Leave No Trace. We would like to see someone use our research to further explore the culture of this small group of hikers and understand why they behave this way and why they have little regard for taking care of the trail.

Such research could lead to activism for hikers to be more environmentally aware. Chancellor et al. argued that “norms matter” in understanding how online health communities promote behavior changes [10]. We suggest that norms could also be integral to infiltrating the community of hikers who are “immune to the Leave No Trace concept and education”. If sufficient data can be found to study this community, efforts can be made to approach this community in a way that can make them **want** to practice Leave No Trace.

It would also be beneficial to group the blog posts together based on which trail they describe, and then run topic analysis on the Triple Crown trails. We could then compare the blog posts' topics to the tweets' topics. Similarly, it would be interesting to see what the tweet topics looked like if the tweets were further separated into clusters of conservancy groups, hikers, news accounts, and outfitters (e.g., REI Co-op, Patagonia, the North Face).

Figure 6.1 shows more tweets than the heat maps in Figures 5.9 and 5.10. In the future, it would be nice to make an effort of harvesting tweets from other major trails around the world to compare trail cultures around the world from an even broader perspective.

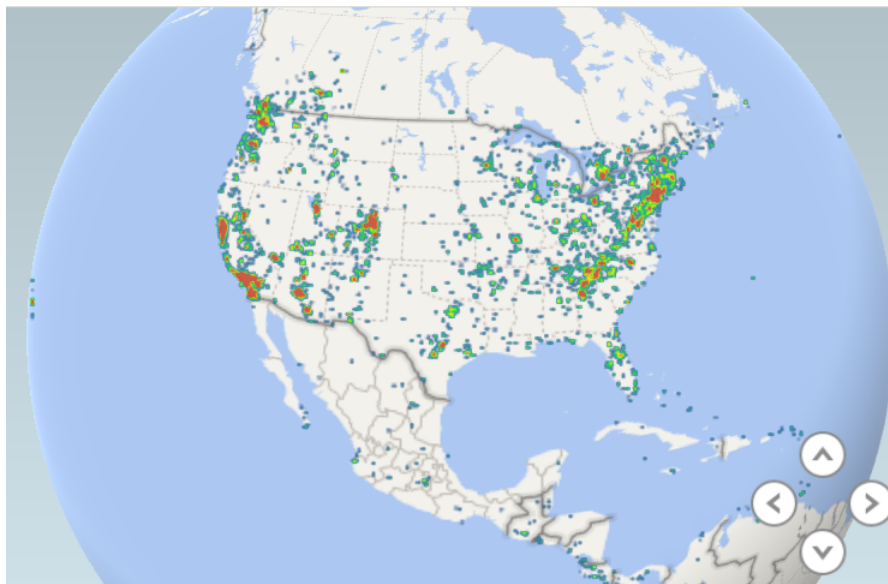


Figure 6.1: Heatmap of tweets in North America [21]

Chapter 7

Summary

This thesis explored a variety of trail cultures (i.e., backpacking culture, conservation culture, culture of the PCT, culture of the AT, culture of trail depreciation) in an effort to understand the values and motivations of trail stakeholders. We were able to understand the cultures of the Triple Crown trails and immerse ourselves in backpacking culture and conservation culture by performing preliminary ethnographic studies. We used Latent Dirichlet Allocation to identify the topics that were valued on the Triple Crown trails and found that the topics of the Appalachian Trail community reflected valuing in-the-moment experiences, whereas the topics of the Pacific Crest Trail community reflected an emphasis on the meticulous planning involved in a long distance hike. With regard to the Continental Divide Trail, we found that there were not enough tweets to reflect the values of the trail. Since the Continental Divide Trail is the newest of the Triple Crown trails, and because the official trail has not yet been established, this was not entirely unexpected; so few hikers attempt to thru-hike the CDT every year that the majority of the tweets in the collection came from its conservancy organization. In building a classifier to identify depreciative behavior, we read through tweets in search of training data, and determined that littering and unleashed pets were by far the biggest concerns to the hiking community on Twitter, but we found that the concerns for trail depreciation were reflected differently in blogs.

Bibliography

- [1] Pacific Crest Trail Association. Pacific Crest Trail Association 2014—2017 strategic plan, July 2015. URL <https://www.pcta.org/wp-content/uploads/2013/10/PCTA-Strategic-Plan-2014-to-2017.pdf?x65851>. (accessed April 12, 2018).
- [2] Pacific Crest Trail Association. Our mission, vision and values, 2018. URL <https://www.pcta.org/about-us/our-mission-vision-and-values/>. (accessed April 12, 2018).
- [3] Abigail Bartolome, Md Islam, and Soumya Vundekode. Clustering and topic analysis in CS 5604 Information Retrieval Fall 2016. 2016. URL <http://hdl.handle.net/10919/73712>.
- [4] Abigail Bartolome, Edward A. Fox, and D. Scott McCrickard. Understanding trail cultures through various stakeholders of the trail. In *ACM CHI 2018 workshop on HCI Outdoors*, Montreal, Canada, 2018.
- [5] Abigail Bartolome, D. Scott McCrickard, and Edward A. Fox. Exploring cultural differences in the Triple Crown trails. In *ACM GROUP 2018 workshop on Technology on the Trail*, Sanibel Island, FL USA, January 2018.
- [6] K. Berger. Pondering the Triple Crown trails: A thru-hiker tries to figure out her favorite. <https://sectionhiker.com/>, 2013. (accessed January 23, 2017).
- [7] Karen Berger. *America’s Great Hiking Trails*. Rizzoli, New York, NY, 2014.
- [8] David M. Blei, Andrew Y. Ng, and Michael I. Jordan. Latent Dirichlet Allocation. *J.*

- Mach. Learn. Res.*, 3:993–1022, March 2003. ISSN 1532-4435. URL <http://dl.acm.org/citation.cfm?id=944919.944937>. (accessed September 2016).
- [9] Matthew Bock. A Framework for Hadoop Based Digital Libraries of Tweets. Master’s thesis, Virginia Tech, Blacksburg, VA, 2017. URL <http://hdl.handle.net/10919/78351> (Accessed June 26, 2018).
- [10] Stevie Chancellor, Andrea Hu, and Munmun De Choudhury. Norms matter: Contrasting social support around behavior change in online weight loss communities. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI ’18, pages 666:1–666:14, New York, NY, USA, 2018. ACM. ISBN 978-1-4503-5620-6. doi: 10.1145/3173574.3174240.
- [11] Jieun Chon, Haitao Wang, Yali Bian, and Shuo Niu. CS5604: Information and Storage Retrieval Fall 2017-FE (Front-end team). 2017. URL <http://hdl.handle.net/10919/81423>.
- [12] Continental Divide Trail Coalition. CDTC background, 2018. URL <http://continentaldivide-trail.org/about-cdtc/cdtc-background/>. (accessed April 12, 2018).
- [13] Appalachian Trail Conservancy. 2017 thru-hiker Registration Charts, 2018. URL <http://www.appalachiantrail.org/home/explore-the-trail/thru-hiking/voluntary-thru-hiker-registration/2016-thru-hiker-registration-charts>. (accessed April 12, 2018).
- [14] Appalachian Trail Conservancy. Our mission, vision, and values, 2018. URL <http://www.appalachiantrail.org/home/about-us/mission-vision-values>. (accessed April 12, 2018).

- [15] Appalachian Trail Conservancy. Thru-Hiking, 2018. URL <http://www.appalachiantrail.org/home/explore-the-trail/thru-hiking>. (accessed April 12, 2018).
- [16] Appalachian Trail Conservancy. 2016 SEES RISE IN ETHICAL HIKING PRACTICES AMONG APPALACHIAN TRAIL HIKERS, 2018. URL <http://www.appalachiantrail.org/home/community/blog/ATFootpath/2016/08/11/2016-sees-rise-in-ethical-hiking-practices-among-appalachian-trail-hikers>. (accessed April 12, 2018).
- [17] Dan Cosley, Dan Frankowski, Sara Kiesler, Loren Terveen, and John Riedl. How oversight improves member-maintained communities. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '05, pages 11–20, New York, NY, USA, 2005. ACM. ISBN 1-58113-998-5. doi: 10.1145/1054972.1054975. URL <http://doi.acm.org/10.1145/1054972.1054975>.
- [18] Cristian Danescu-Niculescu-Mizil, Robert West, Dan Jurafsky, Jure Leskovec, and Christopher Potts. No country for old members: User lifecycle and linguistic change in online communities. In *Proceedings of the 22Nd International Conference on World Wide Web*, WWW '13, pages 307–318, New York, NY, USA, 2013. ACM. ISBN 978-1-4503-2035-1. doi: 10.1145/2488388.2488416. URL <http://doi.acm.org/10.1145/2488388.2488416>.
- [19] Alan Dix. Mental geography, wonky maps and a long way ahead. In *GeoHCI, Workshop on Geography and HCI*, CHI, 2013. URL <http://alandix.com/academic/papers/GeoHCI2013/>.
- [20] Andrew D. Eason, Kevin M. Cianfarini, Marshall C. Hansen, and Shane J. Davies. Blog and forum collection for trail study. *CS4624 Multimedia, Hypertext, and Information*

- Access Final Report*, 2018. URL <http://hdl.handle.net/10919/83214>. Blacksburg, VA.
- [21] Mark Episcopo, Vedant Tyagi, Shivani Rajasekaran, and Patricio Moreno. Twitter tweet visualization on trails. *CS4784 Technology on the Trail Capstone Final Report*, 2018. Blacksburg, VA.
- [22] Anatoliy Gruzd, Barry Wellman, and Yuri Takhteyev. Imagining twitter as an imagined community. *American Behavioral Scientist*, 55(10):1294–1318, 2011. doi: 10.1177/0002764211409378. URL <https://doi.org/10.1177/0002764211409378>.
- [23] Ellie Harmon and Melissa Mazmanian. Stories of the smartphone in everyday discourse: Conflict, tension and instability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 1051–1060, New York, NY, USA, 2013. ACM. ISBN 978-1-4503-1899-0. doi: 10.1145/2470654.2466134. URL <http://doi.acm.org/10.1145/2470654.2466134>.
- [24] Mary E Harmon. *Computing as Context: Experiences of Dis/Connection Beyond the Moment of Non/Use*. PhD thesis, University of California, Irvine, 2015.
- [25] Steve Harrison. Research through design. Graduate Special Topics Course, 2018. Virginia Polytechnic Institute and State University, Blacksburg, VA.
- [26] Joe Hikes. “Hike Your Own Hike”: What does it mean?, Jan 2015. URL <http://joehikes.tumblr.com/post/107841653149/hike-your-own-hike-what-does-it-mean>. Accessed June 26, 2018. Linked from the Journalist community on the PCTA website <https://www.pcta.org/journalist/post/hike-your-own-hike-what-does-it-mean/>.

- [27] hikinghammonds. About the CDT and the Triple Crown. 2017. URL <https://hikinghammonds.com/2017/12/04/about-the-cdt-and-the-triple-crown/>.
- [28] Zunaira Jamil, Diana Inkpen, Prasadith Buddhitha, and Kenton White. Monitoring Tweets for Depression to Detect At-risk Users, 01 2017.
- [29] Hanlin Li, Lynn Dombrowski, and Erin Brady. Working Toward Empowering a Community: How Immigrant-Focused Nonprofit Organizations Use Twitter During Political Conflicts. In *Proceedings of the 2018 ACM Conference on Supporting Groupwork*, GROUP '18, pages 335–346, New York, NY, USA, 2018. ACM. ISBN 978-1-4503-5562-9. doi: 10.1145/3148330.3148336. URL <http://doi.acm.org/10.1145/3148330.3148336>.
- [30] Jeffrey Marion. *Leave No Trace in the Outdoors*. Stackpole Books, Mechanicsburg, PA, 2014.
- [31] Jeffrey Marion. On groups immune to LNT education. personal conversation, January 2018.
- [32] Jeffrey L. Marion. Less typical now: A 1970s retrospective. *Appalachia*, Winter/Spring: 8–19, 2007.
- [33] Jeffrey L. Marion, Yu-Fai Leung, Holly Eagleston, and Kaitlin Burroughs. A review and synthesis of recreation ecology research findings on visitor impacts to wilderness and protected natural areas. *Journal of Forestry*, 114(3):352–362, 2016. doi: 10.5849/jof.15-498. URL <http://dx.doi.org/10.5849/jof.15-498>.
- [34] Jaya Marr. Be as fit as a mountain guide. Email Newsletter, 2016. The Mountain Training School.

- [35] Michele Mauri, Tommaso Elli, Giorgio Caviglia, Giorgio Uboldi, and Matteo Azzi. RAWGraphs: A Visualisation Platform to Create Open Outputs. In *Proceedings of the 12th Biannual Conference on Italian SIGCHI Chapter, CHIItaly '17*, pages 28:1–28:5, New York, NY, USA, 2017. ACM. ISBN 978-1-4503-5237-6. doi: 10.1145/3125571.3125585. URL <http://doi.acm.org/10.1145/3125571.3125585>.
- [36] D. Scott McCrickard. Technology on the trail. Capstone Course, 2018. Virginia Polytechnic Institute and State University, Blacksburg, VA.
- [37] S. Mehta and Vinayagam R. CS5604: Extracting topics from tweets and webpages for IDEAL, May 2016. URL <http://hdl.handle.net/10919/70933>. (accessed August 2016).
- [38] Ankan Mullick, D SurjodoyGhosh, Shivam Maheswari, Srotaswini Sahoo, Suman Kalyan Maity, C. Soumya, and Pawan Goyal. Identifying opinion and fact subcategories from the social web. In *GROUP*, 2018.
- [39] Bonnie A. Nardi, Diane J. Schiano, and Michelle Gumbrecht. Blogging as Social Activity, or, Would You Let 900 Million People Read Your Diary? In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work, CSCW '04*, pages 222–231, New York, NY, USA, 2004. ACM. ISBN 1-58113-810-5. doi: 10.1145/1031607.1031643. URL <http://doi.acm.org/10.1145/1031607.1031643>.
- [40] Carman Neustaedter and Phoebe Sengers. Autobiographical design in HCI research: Designing and learning through use-it-yourself. In *Proceedings of the Designing Interactive Systems Conference, DIS '12*, pages 514–523, New York, NY, USA, 2012. ACM. ISBN 978-1-4503-1210-3. doi: 10.1145/2317956.2318034. URL <http://doi.acm.org/10.1145/2317956.2318034>.
- [41] ATC Board of Managers. The Appalachian Trail experience. July/August 1997.

- [42] Roanoke Outside Foundation. Appalachian Trail hikes. <http://www.roanokeoutside.com/land/hiking/appalachian-trail-hikes/>. accessed May 28, 2018.
- [43] National Park Service. Depreciative visitor use, December 2017. URL https://www.nps.gov/shen/learn/nature/depreciative_visitor_use.htm. (accessed March 2018).
- [44] Norman Makoto Su and EunJeong Cheon. Reconsidering nature: The dialectics of fair chase in the practices of American midwest hunters. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 6089–6100, New York, NY, USA, 2017. ACM. ISBN 978-1-4503-4655-9. doi: 10.1145/3025453.3025966. URL <http://doi.acm.org/10.1145/3025453.3025966>.
- [45] C. Zhai and S. Massung. *Text Data Management and Analysis*. Association for Computing Machinery and Morgan & Claypool Publishers, San Francisco, CA, 2016.
- [46] Dejin Zhao and Mary Beth Rosson. How and why people Twitter: The role that microblogging plays in informal communication at work. In *Proceedings of the ACM 2009 International Conference on Supporting Group Work*, GROUP '09, pages 243–252, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-500-0. doi: 10.1145/1531674.1531710. URL <http://doi.acm.org/10.1145/1531674.1531710>.

Appendices

Appendix A

Map of Triple Crown Trails

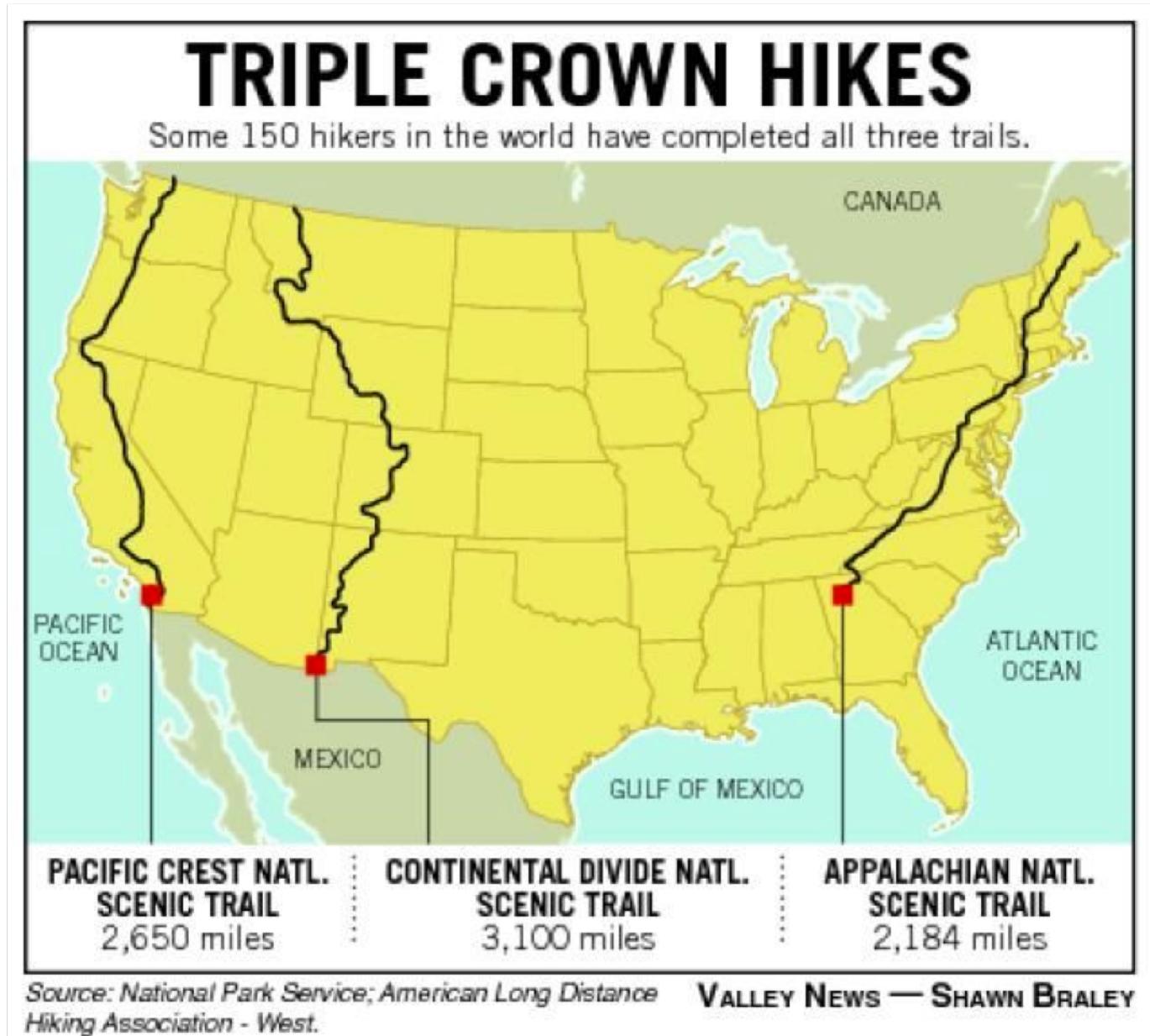


Figure A.1: Map of Triple Crown Hikes [27]

Appendix B

ACM GROUP 2018 Technology on the Trail Workshop

Position Paper [\[5\]](#)

Exploring Cultural Differences in the Triple Crown Trails

Abigail Bartolome

Dept. of Computer
Science
Virginia Tech
abijbart@vt.edu

D. Scott McCrickard

Dept. of Computer
Science
Virginia Tech
dmccrick@cs.vt.edu

Edward A. Fox

Dept. of Computer
Science
Virginia Tech
fox@vt.edu

GROUP 2018 Position Paper

In an essay on *SectionHiker.com*, Karen Berger reflects on her experiences on the Triple Crown Trails (Appalachian Trail, Pacific Crest Trail, Continental Divide Trail) in an attempt to answer a popular question that many Triple Crown veterans are asked—which one is your favorite? [1]. As she sorts through her memories of her thru hikes, she muses, “each one is different; each has a soul”. My research takes inspiration from this musing. What makes these trails different from each other—aside from their physical differences (e.g., trail length, geographical location, wildlife, etc.?). Can we describe the “souls” of these trails?

If there is one thing that anyone needs to know about me, it is that I am *very* enthusiastic about outdoor adventures—primarily in regards to hiking. As I have gotten older and have experienced more hikes, I have grown aware of the cultural differences of different trails as well as the different personas of hikers on these trails. About two years ago, I had a conversation with a ridge runner on the Appalachian Trail in Great Smoky Mountain National Park that resonates with me to this day. She told me that she wrote her undergraduate thesis comparing the Appalachian Trail to a religion. Of course, this in itself was a very interesting conjecture, but what she said after that was

very inspiring. She said that she studied the mission statement and vision of the Appalachian Trail, as well as testimonials and other writings of the hikers, and found that the language of the mission statement and writings were much more spiritual than those of the Pacific Crest Trail and Continental Divide Trail.

This casual conversation on my backpacking trip was incredibly impactful, and it inspired my research efforts for the following two years. This ridge runner had a point. Not all trails are the same, including the Triple Crown Trails. What is it about these trails that makes them each so different? The Appalachian Trail, Pacific Crest Trail, and Continental Divide Trail are all national scenic trails whose mileages are well over 2,000 miles. Needless to say, hikers that embark on hiking these trails in one straight push must be particularly ambitious and have some sort of connection to the great outdoors. But there must be some difference between hikers that decide to take on the Appalachian Trail versus hikers that take on the Pacific Crest Trail—it cannot just be a matter of geographical convenience, right? And even if the trail that a hiker chooses is based on geographical convenience, hikers on each of the respective trails must have very different experiences. Consider Dr. Ellie Harmon’s source of maps on her Appalachian Trail thru hike versus her Pacific Crest Trail hike. When she hiked the Appalachian Trail, like many, all she needed was a guidebook as her source of direction. However, when she hiked the Pacific Crest Trail, she required her phone as her guide. She wrote in



Figure 1 Danner hiking boot, from the Danner retail site

her dissertation that her phone was her *everything* on her Pacific Crest Trail thru hike [2]. I wanted to dedicate my research to learning what makes each of these trails different. What are the cultural differences between the trails? Do the general community of hikers on each of these trails have different values?

I saw an opportunity to explore these questions since my lab (the Digital Library Research Laboratory) is continuously collecting tweets. I began a collection of tweets relating to the Triple Crown trails. I collected tweets that discussed the Appalachian Trail, the Pacific Crest Trail, and the Continental Divide Trail, as well as

tweets written by the conservancy organizations. After all, one would expect that a tweet from the Appalachian Trail Conservancy would reflect the values of the Appalachian Trail community. By using Natural Language Processing techniques, I can learn what these communities care about.

During my early tweet analysis, I was able to determine some trends in the topics for each of the trail communities (Table 1). The tweet collection for these specific trails began in January 2017. By May 2017, the tweets started to reveal unique topics specific to each trail. The tweets about the Pacific Crest Trail were very

	Appalachian Trail Topics	Pacific Crest Trail Topics	Continental Divide Trail Topics
Topic 1	#indigenous, #tairp, #amerianindian8, day, knob, mcafee, trailva	California, #pct2017, 2, story, tips, resupply, #pics	help, #bravethecdt, #hikecdt, today, @cdnst1, vote, great
Topic 2	va, catawba, sunrise, halfway, #backpacking, just, oc4444x2400	@pctassociation, like, today, #lwcf, win, did,great	#bravethecdt, @cdnst1, #hikecdt, help, great, support, today
Topic 3	days, amp, long, mountain, complete, miles, week	mount, adams, goat, rocks, @hogansog, washington, view	#bravethecdt, help, today, @cdnst1, support, vote, great
Topic 4	#travel, #bestseller, black, 1, awol, books, 2	#orshow, booth, gear, come, free, #pct2017, @danner	#bravethecdt, help, today, great, @cdnst1, support, #hikecdt
Topic 5	new, going, woman, 80yearold, solo, sisters, twin	wild, lost, #travel, #bestseller, oprahs, #7, #8	#bravethecdt, #hikecdt, @cdnst, support, help, great, today
Topic 6	hiker, #at2017, @thetrekat, 5, update, thruhiker, thruhikers	taking, #backpacking, months, job, better, 4, day	#hikecdt, #bravethecdt, @cdnst1, great, help, support, today
Topic 7	#at2017, @thetrekat, #trail, gear, list, things, #photography	#pctdays, new, instagram, year, weeks, posts, bitesized	help, #bravethecdt, today, #votecdt, vote, 25k, @cdnst1

Table 1 Results of early topic analysis

related to logistics and planning hikes. One of the topics was resupply tips in California, another topic was about acquiring Danner hiking boots (Figure 1) at an Outdoor Research show. Danner is renowned for making durable hiking boots, and Outdoor Research is a very reputable company specializing in outdoor gear. There was also a Pacific Crest Trail topic about hikers taking months off from their jobs to complete their thru hike. Since northbound hikers will typically start their thru hikes between mid-April and early May, it makes sense that the tweets written in the months leading to the peak weeks for beginning a Pacific Crest Trail thru hike would be about planning the logistics of a thru hike (i.e., where to resupply in California, where to acquire gear, taking time off from work).

The Appalachian Trail community's tweets were more reflective of in-the-moment experience on the trail. One topic referred to McAfee Knob—an iconic lookout point on the Appalachian Trail. There was a topic that was about watching the sunrise on a backpacking trip in the Catawba Valley in Virginia (the Appalachian Trail runs through the Catawba Valley), and another topic describes a woman's hiking trip. There could be a variety of reasons for the difference in the topic themes between the Appalachian Trail communities and the Pacific Crest Trail communities. Appalachian Trail thru hikes usually begin around March, so by May it was possible that a majority of the tweets in the collection were made by hikers who had already begun their thru hikes, whereas many of the Pacific Crest Trail hikers were just starting. Another possibility is that there is a higher concentration of Pacific Crest Trail hikers who would tweet about thru hike planning. On the other hand, while there are many hikers who attempt to thru hike and section hike the Appalachian Trail, the Appalachian Trail is also popular to casual day hikers. It is possible that there are just a lot of day hikers tweeting about their day on the trail.

Understanding the differences between the hikers on these trails, as well as understanding the difference in culture and values of the trail communities, is an aim of my research. We are nearing a year of collecting data on the Triple Crown trails, and I would be interested in doing a month-by-month study on the topics found in each trail to see if there is a general model for topics that are discussed by backpackers over the course of a year.

The "Identifying Roles" and "Channeling Conflicts" workshop activities are particularly interesting to me because I have an interest in identifying who is part of each of these trail communities. For example, I know that the Appalachian Trail community is made of thru hikers and day hikers, as well as park rangers who enforce Leave No Trace practices. However, I want to have a more defined understanding of the roles played by the members of these trail communities. Similarly, I want to understand the ideologies of the members that comprise these trail communities, and I believe that the "Channeling Conflicts" activity will provide a way to understand the ideologies of these communities. I also feel that I could contribute my own hiking ideologies, as an avid hiker, as well as the ideologies that I can infer from my own research, to this workshop activity. Since I have been on a few backpacking trips, and many more hikes, I also feel that I could contribute my feelings of using technology on the trail to the "Brainstorming Beginnings" workshop.

Acknowledgements

I would like to thank and acknowledge Dr. Scott McCrickard for giving me the opportunity to work on a project that was so closely aligned with my research interests, and that I could enthusiastically put my efforts toward. I would like to acknowledge the Digital Library Research Laboratory, especially those associated with the Integrated Digital Event Archiving and Library (IDEAL) and Global Event and Trend

Archive Research (GETAR) projects. Thus, I would also like to express my appreciation to the National Science Foundation for funding the IDEAL and GETAR projects (grants IIS-1319578 and IIS-16119028).

References

[1] Berger, K. (2016, May 31). Pondering the Triple Crown Trails: A Thru-hiker tries to Figure out her Favorite. In SectionHiker.com. Retrieved January 23, 2017.

[2] Harmon, Mary E. "Computing as Context: Experiences of Dis/Connection Beyond the Moment of Non/Use." (2015).

Appendix C

ACM CHI 2018 HCI Outdoors Workshop

Position Paper [\[4\]](#)

Understanding Trail Cultures through Various Stakeholders of the Trail

Abigail Bartolome
Virginia Tech
Blacksburg, VA 24061 USA
abijbart@vt.edu

Edward A. Fox
Virginia Tech
Blacksburg, VA 24061 USA
fox@vt.edu

Scott McCrickard
Blacksburg, VA 24061 USA
mccricks@cs.vt.edu

Authors maintain the copyright. Workshop on HCI Outdoors: Understanding Human-Computer Interaction in the Outdoors at CHI 2018, April 21, 2018, Montreal, Canada

Abstract

In this position paper, we explore the cultural issues associated with HCI outdoors. If we describe an enthusiasm for trail life and the outdoors as "trail culture", then we need to be aware of the different subcultures within trail culture. The hiking community has a variety of stakeholders who have different values and motivations for being on the trail. Each grouping of stakeholders has their own culture. These cultures should be explored to understand if or how HCI should be used on trails. In this paper, we provide examples of different trail cultures and ways that trail cultures can be studied.

Author Keywords

Culture; social media; social computing; community

ACM Classification Keywords

J.4 [Social and Behavioral Sciences]: Computer Applications.

Introduction

Triple Crown veteran Karen Berger wrote an essay on *SectionHiker.com* reflecting on her thru-hiking experiences on the Appalachian Trail, Pacific Crest Trail, and Continental Divide Trail and mused "each one is different; each has a soul" [2]. As she compared and contrasted each trail, it became evident that the differences between

the trails went beyond the physical differences (e.g., trail mileage, geographical location, wildlife, etc.). The differences seemed to be cultural. It seemed that there were subcultures within the culture of trail life. In an earlier study, which will be further explained in the Data Analysis section, there was a definitive difference between the Appalachian Trail and the Pacific Crest Trail. To further pursue the study of cultural differences among the trails, we are looking at the cultures of different stakeholders on the trails.

Within the trail community, there are a variety of stakeholders who have different motivations for being on the trail, and who value different aspects of trail life. At the ACM GROUP 2018 Technology on the Trail workshop, Michael Horning, Professor in Virginia Tech's Department of Communication, and Lindah Kotut, a Ph.D. student in Computer Science at Virginia Tech, held an activity in which participants clustered different stakeholders of technology on the trail (e.g., thru-hikers, section hikers, day hikers, trail angels, ridge runners, park rangers, farmers, families, etc.) and examined the similarities and differences in their values and goals [3]. Do these stakeholders have different trail cultures that lead to their differences in values and goals? If so, these cultures should be studied and considered when designing HCI to be used on the trail.



Figure 1: Jeffrey Marion on a reunion thru-hike on the Appalachian Trail [4]

Examples from Backpacking Culture

When many people imagine the stereotypical hiker, they may picture a hiker with days worth of gear strapped to their back, such as the hikers depicted in Figure 1. These hikers are called backpackers. Some backpackers may go out for a few days at a time (section hikers), while the more devoted may go out for several months at a time to complete a thru-hike. Whether a backpacker is out for a

50 mile hike over a 3 day period or spending months thru-hiking all 2,190 miles of the Appalachian Trail, backpackers are a devoted group of hikers with their own unique culture.

In a reflective article comparing modern backpacking to backpacking in the 1970's, Jeffrey Marion writes, "No backpacker with a few thousand miles under his feet fails to consider the weight of every packed item. The ultra-light paradigm places greater focus on a new array of light-weight gear, your knowledge and skills, and doing without" [4]. Managing trail weight is a common culture in backpacking. There are extreme practices like shortening the handle of a toothbrush or repackaging food to lessen the weight of the backpack. Such practices seem extreme to those outside of the backpacking community. The term "comfort item" is often used by backpackers to refer to items considered "non-essential," that make the backpack heavier than it needs to be. However, when backpackers discuss bringing a 1-pound camping chair or a warmer, fluffier (i.e., heavier) sleeping bag because it makes their trail experience more enjoyable, the untrained ear may struggle to understand how the difference between a 29-pound backpack and a 30-pound backpack could cause such a disturbance to a physically fit backpacker. In the backpacking community, the notion of spending an extra \$100 on a smaller tent is justifiable if it means reducing trail weight. Figure 2 shows a comparison of two backpacking tents. Both are 2-person, 3-season, and made by Big Agnes. The main difference between the Fly Creek and the Frying Pan is that the Fly Creek weighs about half as much as the Frying Pan, and accordingly costs an extra \$140. Oftentimes, in backpacking culture, saving the 2+ pounds in trail weight is well worth the \$140.

Reviews	★★★★	★★★★	hide
Best Use	Backpacking	Backpacking	hide
Seasons	3-season	3-season	hide
Sleeping Capacity	2-person	2-person	hide
Minimum Trail Weight	1 lb. 15 oz.	4 lbs. 5 oz.	hide
Packaged Weight	2 lbs. 5 oz.	5 lbs. 5 oz.	hide
Packed Size	4 x 19 inches	18 x 7 inches	hide

Figure 2: Comparing two 2-person backpacking tents by Big Agnes on REI.com

Examples from Conservation Culture

The outdoor spaces that are so loved and coveted by trail enthusiasts are protected by hikers and conservancy organizations. One well-known conservancy organization is the Leave No Trace Center for Outdoor Ethics. Practicing Leave No Trace is common in the conservation culture. Practices as common as “pack-it in, pack-it out” are part of Leave No Trace. Other less visible practices such as digging cat-holes (holes for solid human waste that are dug 4-6 inches in diameter and 6-8 inches in depth) 200 feet from the trail and water sources, or minimizing visible impact after camping in pristine areas, also are part of Leave No Trace, and thus conservation culture [1].

Data Analysis

We have made efforts toward understanding the different trail cultures through data analysis. From January 2017 to January 2018, we collected over 1.5 million tweets from

trail enthusiasts. Some tweets are by backpackers, day hikers, conservancy organizations, and outdoor outfitters. In May of 2017, we separated the tweets based on which Triple Crown Trail they were about. Figure 3 shows results from topic analysis of our data using LDA.

The tweets about the Pacific Crest Trail were very related to logistics and planning hikes. One of the topics was resupply tips in California; another topic was about acquiring Danner hiking boots (Figure 1) at an Outdoor Research show – Danner is renowned for making durable hiking boots, and Outdoor Research is a very reputable company specializing in outdoor gear. There was also a Pacific Crest Trail topic about hikers taking months off from their jobs to complete their thru-hike. Since northbound hikers will typically start their thru-hikes between mid-April and early May, it makes sense that the tweets written in the months leading to the peak weeks for beginning a Pacific Crest Trail thru-hike would be about planning the logistics of a thru-hike (i.e., where to resupply in California, where to acquire gear, taking time off from work).

The Appalachian Trail community’s tweets were more reflective of in-the-moment experience on the trail. One topic referred to McAfee Knob — an iconic lookout point on the Appalachian Trail. There was a topic about watching the sunrise on an Appalachian Trail backpacking trip in the Catawba Valley in Virginia, and another topic describes a woman’s hiking trip. There could be a variety of reasons for the difference in the topic themes between the Appalachian Trail communities and the Pacific Crest Trail communities. Appalachian Trail thru-hikes usually begin around March; by May, it is possible that a majority of the tweets were made by hikers who had already begun their thru-hikes, whereas many of the Pacific Crest Trail

hikers were just starting. Another possibility is that there is a higher concentration of Pacific Crest Trail hikers who would tweet about thru-hike planning. On the other hand, while there are many hikers who attempt to thru-hike and section hike the Appalachian Trail, the Appalachian Trail is also popular among casual day hikers. It is possible that there are just a lot of day hikers tweeting about their day on the trail.

Understanding the differences between the hikers on these trails, and understanding the differences in culture and values of the Triple Crown Trail communities, are aims of the research. Further, now that we have almost twelve months of data on the Triple Crown trails, it is becoming feasible to launch a month-by-month study on the topics found in each trail to devise a seasonal model for topics that are discussed by backpackers.

Future Work

From speaking to forestry specialists and studying samples of tweets from our dataset, we have found less traditional examples of trail culture, that we plan to study. Jeffrey Marion, founding member of the Leave No Trace board of directors and research biologist, expressed an interest in understanding the small group of hikers who actively avoid Leave No Trace practices — hikers who refuse to pack out their trash, refuse to dig cat-holes, and insist on burning their leftover food. Understanding their motivations to ignore conservation practices could aid in educating them on the importance of conserving the land that they are enjoying during their hike.

Studying tweet samples also led us to a large number of tweets about nudity on the trail. These tweets had words and phrases like "nude", "naked", and "as nature intended". There has been an interest in understanding

why people want to hike in the nude, and furthermore, why they publicly talk about hiking in the nude.

We expect to find other topics and issues, and to integrate our findings from interviews with those from data analysis, to develop a deeper understanding of trail cultures.

	Appalachian Trail Topics	Pacific Crest Trail Topics	Continental Divide Trail Topics
Topic 1	#indigenous, #taip, #ameraindians, day, knob, micafse, trailva	California, #pct2017, 2, story, tips, resupply, #pics	help, #bravethecdt, #hikecdt, today, @cdnst1, vote, great
Topic 2	va, catabwa, sunrise, halfway, #backpacking, just, cc444x2400	@pctassociation, like, today, #fwd, win, did, great	#bravethecdt, @cdnst1, #hikecdt, help, great, support, today
Topic 3	days, amp, long, mountain, complete, miles, week	mount, adams, goat, rocks, @hegansog, washington, view	#bravethecdt, help, today, @cdnst1, support, vote, great
Topic 4	#travel, #bestseller, black, 1, #wol, books, 2	#rshow, booth, gear, come, free, #pct2017, @damner	#bravethecdt, help, today, great, @cdnst1, support, #hikecdt
Topic 5	new, going, woman, 80yearold, solo, sisters, twin	wild, lost, #travel, #bestseller, oprahs, #7, #8	#bravethecdt, #hikecdt, @cdnst, support, help, great, today
Topic 6	hiker, #at2017, @thetrekat, 5, update, thruhiker, thruhikers	taking, #backpacking, months, job, better, 4, day	#hikecdt, #bravethecdt, @cdnst1, great, help, support, today
Topic 7	#at2017, @thetrekat, #trail, gear, list, things, #photography	#pctdays, new, instagram, year, weeks, posts, blessed	help, #bravethecdt, today, #votecdt, vote, 25k, @cdnst1

Figure 3: Results of early topic analysis.

Acknowledgements

I would like to thank and acknowledge Dr. Edward Fox and Dr. Scott McCrickard for giving me the opportunity to work on a project that is so closely aligned with my interests, and upon which I could enthusiastically focus. I also would like to thank the CS 4784 Technology on the Trail and CS4624 Multimedia, Hypertext, and Information Access capstone classes for their contributions to this research effort. I acknowledge the Digital Library Research Laboratory, especially those associated with the Integrated Digital Event Archiving and Library (IDEAL) and Global Event and Trend Archive Research (GETAR) projects. Thus, I also express my appreciation to the National Science Foundation for funding the IDEAL and GETAR projects (grants IIS-1319578 and IIS-16119028).

References

- [1] *Leave No Trace in the Outdoors*. Stackpole Books, Mechanicsburg, PA, 2014.
- [2] Berger, K. Pondering the triple crown trails: A thru-hiker tries to figure out her favorite. <https://sectionhiker.com/>, 2013 (accessed January 23, 2017).
- [3] Horning, M., and Kotut, L. Who's out there? what are they doing? In *ACM GROUP 2018 workshop on Technology on the Trail*. (Sanibel Island, FL USA, Jan. 2018).
- [4] Marion, J. L. Less typical now: A 1970s retrospective. *Appalachia Winter/Spring* (2007), 8–19.