

© 2019 by Xueqing Liu. All rights reserved.

ASSISTING THE COMMUNICATION BETWEEN END USERS AND MOBILE
SYSTEMS WITH NATURAL LANGUAGE INTERFACES

BY
XUEQING LIU

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Computer Science
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2019

Urbana, Illinois

Doctoral Committee:

Professor ChengXiang Zhai, Chair
Professor Tao Xie, Director of Dissertation
Professor Carl Gunter
Associate Professor William Enck

Abstract

Mobile devices are ubiquitous in people’s lives. As mobile users are expected to exceed 4.5 billion in 2019, mobile devices are the closest approach for users to communicate with the outside world. It is important to ensure such communication goes smoothly. That is, on the one hand, users can easily and conveniently express their need to the system, and on the other hand, users should be able to understand and trust the requests and decisions from the system.

Same as in other software platforms, poor communications between users and mobile devices can result in the failure in requirements understanding, reduction of system reliability and lower user satisfaction. However, the communication on mobile devices is especially challenging due to three unique characteristics of mobile devices. First, it is more difficult to perform typing on mobile devices, also mobile users highly rely on voice input [voi, 2018, goo, 2016], the two facts limit the types of information that users can conveniently input on mobile devices (e.g., technical terms, programming language). Second, mobile screens are smaller, which limits the amount of information that users can focus on at the same time. Third, the challenges in mobile security and privacy directly caused the challenge in user trust, transparency issue and communication problems regarding how user privacy is protected over mobile devices [loc, 2018].

This thesis seeks to provide a holistic picture on the new communication challenges brought by mobile devices. More specifically, it studies the communication through natural language because the natural language is the most common way for end users to communicate. This thesis is organized as follows. Chapter 1 introduces the background of formal communication models, then introduces the unique communication problem on mobile devices. The three challenges above are discussed from Chapter 2 to Chapter 4. For each challenge, we first provide an overview of the state of the art in existing literature, or perform empirical study to gain understanding on questions that are yet to be answered. Then, based on the analysis results, we identify feasible technical problems, and build solutions to support the communication process. Finally, Chapter 5 summarizes the remaining challenges, discusses the future work and draw the conclusion.

Table of Contents

List of Tables	v
List of Figures	vi
List of Abbreviations	vii
List of Symbols	viii
Chapter 1 Introduction	1
1.1 Why Do Human Communicate?	2
1.2 The Importance of Studying System-User Communication	3
1.3 Shannon-Weaver Communication Model	4
1.4 System-User Communication in Natural Language	6
1.4.1 Natural Language Input Interface	6
1.4.2 Natural Language Output Interface	6
1.5 An Overview of the Mobile Systems and User Interfaces	7
1.6 The Challenges of End User Communications Over Mobile Devices	8
1.6.1 Challenge 1: Making Security Decisions	8
1.6.2 Challenge 2: e-Commerce Shopping	9
1.6.3 Challenge 3: Performing Data Analysis	10
1.7 Thesis Overview	11
Chapter 2 Assisting Security Decision Making with Natural Language Explanations . .	13
2.1 Overview of Android Security Decision Making	15
2.1.1 Android Infrastructure and Third-party Market	15
2.1.2 Android Security System	15
2.1.3 Over-privileged Problem and User Reaction	15
2.1.4 Runtime Permission System	15
2.2 Review Existing Work	15
2.2.1 Static/dynamic Analysis	15
2.2.2 Security Risk Assessment	15
2.2.3 Natural Language Explanation	15
2.3 An Empirical Study on Natural Language Explanations in Existing Apps	15
2.3.1 Motivation of Empirical Study	15
2.3.2 Data Collection Process	15
2.3.3 RQ1: Population	15
2.3.4 RQ2: Categorical Explanation Ratio	15
2.3.5 RQ3: Incorrect Explanations	15
2.3.6 RQ4: Specificity of Explanations	15
2.3.7 RQ5: Comparing App Description and Runtime Rationale	15
2.4 An Analysis of Difficulty of Explaining Permissions	15
2.5 CLAP: A Recommender System for Assisting Apps Explain Permissions	16

2.5.1	Module 1: Similar App Retrieval	16
2.5.2	Module 2: Permission Explanation Detector	16
2.5.3	Module 3: Sentence Splitter	16
2.5.4	Module 4: Sentence Reranker	16
2.5.5	Module 5: Sentence Post-processor	16
2.6	Experiments on CLAP	16
2.6.1	Data Collection	16
2.6.2	Evaluation Metric	16
2.6.3	Baselines	16
2.6.4	Experiment 1: Quantitative Study	16
2.6.5	Experiment 2: Qualitative Study	16
2.7	Summary	16
Chapter 3	Assisting Browsing and Navigation in Search Engine	17
3.1	Existing Solutions for the Small-Screen Browsing	19
3.2	Existing Solutions on Faceted Navigation	19
3.2.1	Faceted Navigation Systems for Structured Search	19
3.2.2	Faceted Navigation Systems for Web Search	19
3.3	Numerical Faceted Search Problem	19
3.3.1	Applications with Numerical Facet Search	19
3.3.2	Evaluation Methodology	19
3.4	Optimizing Numerical Facet Ranges	19
3.4.1	Baseline 1: Quantile	19
3.4.2	Baseline 2: Click Model and Dynamic Programming	19
3.4.3	Parameterizing the Problem	19
3.4.4	Asymptotic Smoothed Objective Function and An Approximately Upper Bound	19
3.4.5	Method 1: Context-Ignoring Method	19
3.4.6	Method 2: Context-Aware Method	19
3.5	Experiments on Numerical Facet Optimization	20
3.5.1	Dataset Collection	20
3.5.2	Observation of Asymptotically Smooth	20
3.5.3	Comparative Results	20
3.5.4	Ablation Study	20
3.6	General Evaluation Challenge for Navigational Systems	20
3.7	A General Framework for Evaluating Navigational Systems	20
3.8	Summary	20
Chapter 4	Assisting Database Queries with Natural Language Interface	21
Chapter 5	Conclusion	22
References	23

List of Tables

List of Figures

1.1	Shannon’s Communication Model [Shannon, 1948]	4
1.2	Weaver’s Communication Model [Shannon et al., 1951]	5
1.3	Due to user knowledge gap and the simplified permission request, the user gets confused on why the app needs permission request	8
1.4	Due to reduced search results and user position bias, the user conduct less exploration of the search results, resulting in the user more likely makes a suboptimal decision	9
1.5	Difference between display space on laptop/desktop vs. on mobile devices	9
1.6	Due to the difficulty to write programs using mobile interface, the user’s need to query the database is expressed in natural language, which leads to the disconnection to the database interface.	10
2.1	Example of Android permission rationale	13
3.1	Amazon’s numerical range interface on mobile device	17

List of Abbreviations

List of Symbols

Chapter 1

Introduction

How is information and knowledge transmitted through communications? Communication is among the most frequent activities in our daily conversations, and it is involved in all forms of knowledge passing, e.g., teachers teaching a course, writers writing a book. In a 2014 book named *a survey on communication study* [Hahn et al., 2014], the authors informally define communication as “...*the process of using symbols to exchange meanings.*” Namely, the meaning must be carried through a list of symbolic choices, i.e., visual or auditory cues. By the time those symbols are selected, it is only selected symbols that will affect the rest of communication, not the original meaning. The most commonly used symbols among humans are verbal languages, or natural languages. However, there is a limitation on the meanings that can be represented by natural language, which will affect the effectiveness of communication.

The academic field of communication studies dates back to ancient Egypt, but one of the foundation work on communication studies in the modern era is the paper by Claude Shannon and Warren Weaver, *The Mathematical Theory of Communications* [Shannon et al., 1951]. This paper proposed a simple 5-element model of communication, which helps to understand the process of knowledge transmission and information flow. Later people pointed out limitations in Shannon-Weaver model, which over-simplifies the human communication process and neglects important factors such as common sense and contexts. As a result, people proposed improved version of this model by fixing those limitations.

In the digital era, communication usually happens over electronic devices. Digital communication means there are more limitations on the symbolic choices of communication. For example, when texting a friend “*I saw a cute dog*”, the only information that is passed on to the friend is the combination of the interface of text message on her phone and the five words. It is thus up to the friend to interpret the meaning of the five words, and the meaning the friend has in mind may or may not be the same as the original meaning. For example, she may be thinking of a different breed than the message sender does.

The communication between user and a computer device highly relies on the physical characteristics of the device and the design of the interface. At the beginning of the 21st century, the communication was changed largely due to the prevalence of world wide web and search engines. Later the mobile devices

dominated the computation market and the communication was changed because of the major difference in mobile interface and system. Today, wearable devices are becoming popular and natural language interfaces are more popular.

This thesis studies the communication process where the communication happens between a human user and a computer system. It seeks to understand the cases when communications are ineffective due to the constraint of computer system, and tries to improve the system to mitigate such limitations. More specifically, this thesis has an emphasize on user communication with mobile devices because mobile devices are the dominant devices now. However, many part of this thesis can also be applied to other devices.

1.1 Why Do Human Communicate?

To give a more comprehensive definition of what is communication, we first take a look at the purposes of communication (so we know that anything that matches the purposes are communications). Here we focus on two-party communications, where the information is passed from the sender (informer) to the receiver (informee). In a 2008 book named “*Information flow and knowledge sharing*”, the authors summarized three universal reasons for communication, whether it is for communication between a human and a system, or communication between two human [Da Silva and Agusti-Cullell, 2008]:

- A . *Communication to explain*: the communication happens when the information sender has some connection to the outside that the informee does not have, and the informee wants to gain such knowledge by directly receiving such information from the informer. The reason may be that informee does have access to the information (e.g., the informee wants to learn the history, she can only learn it from an older person or a text book, i.e., informer), or that it is time or money consuming for the informee to obtain the information by herself (e.g., the informee wants to learn about south pole but does not have time to go there, so she reads the book written by people who went there, i.e., informer).
- B . *Communication to command*: the informer wants to build something to change the outside environment, the informee shares the same goal or has more expertise than the informer, so the informer uses a command to ask the informee to build the task for the informer. For example, a construction worker (informer) wants to build a house, but she does not have the designs of the house, so she seeks a blueprint from a house designer by describing what the house looks like (informee).
- C . *Communication to satisfy*: here the reality of the world is replaced by certain specified requirements to satisfy, and instead of communicating to the reality, the informee accept the alternative reality defined by the set of requirements given by the informer.

The three purposes of communication applies to all software systems that interacts with users. For *explaining*, the software serves as the informer to explain the outside world to the user, e.g., a shopping assistant explains to the user the available items and their prices. For *commanding*, an example is an enterprise database system, or an intelligent assistant such as Amazon Alexa. For *satisfying*, an example is a role playing game where the user plays by following certain rules specified by the game.

1.2 The Importance of Studying System-User Communication

Communication has been an important research topic in the software engineering research community. During the requirements engineering step, stakeholders communicate with each other to better learn the customers' need and transform such needs into concretized requirements. During the development, project managers communicate with developers to assign concretized tasks (communicate to command). At deployment, the software needs to communicate with the user to explain its decisions, e.g., credit card application results, or how is user's data used by the software.

Communication is an important step for enhancing trust and fairness of automated software decision. As artificial intelligence witnesses great success, there has been a huge demand for developing more explainable models for AI software, especially when the automated decisions have an impact on users' personal benefit, such as safety, security, and monetary factors. An example is to help users build trust towards the decisions on driving direction made by autonomous driving vehicles. A recent project by Nvidia explain to the user which parts of the road scene causes the software make what decision on the direction of the steering wheel [exp, 2018]. The General Data Protection Regulation (GDPR) established in May 2018 addressed the importance of transparent decision making with multiple articles/recitals including the "*right to explain*" and "*the principle of transparency*". In particular, Article 22 in GDPR particularly addresses the case of automated decision making, such as when machine learning is involved [aut, 2018]. Under such cases, the users have the right to know why the software/company reaches such decision results.

Another goal of communication concerns the transparency of user data processing. GDPR specifies users' right about the data they own, including location, email address, phone number, etc. Users have the following right on their data: *right to know the access*: users have the right to know what data is collected and know how the data is used, e.g., personalize offering and advertisement; *right to erase*: users have the right to ask their personal data to be erased from the company's server at any time. GDPR also poses high pressure on the compliance of such regulations, the highest fine can cost up to 20 million EUR or 4% of the company's income. To big companies such as Google and Facebook, such fine could mean billions of

dollars. For small companies, they might have to shut down their business. With the advent of GDPR, many companies have updated their privacy policies explaining how user data is processed under their business models [Degeling et al., 2018]. Meanwhile, not only did the companies reacted on their privacy policies, the user interfaces have also been updated which directly communicate with users on the data usage. For example, the mobile user interfaces of Facebook and Twitter have been updated which enable the user to download her processed data directly from the app. The interfaces also detailed the data categories to be opted out.

1.3 Shannon-Weaver Communication Model

The Shannon-Weaver communication model was first introduced by a paper in 1949 named “*A Mathematical Theory of Communication*” by Claude E. Shannon [Shannon, 1948]. The main contributions in this paper includes: it first proposed to quantitatively measure the amount of information as *entropy*; it defined communication as “...*reproducing at one point either exactly or approximately a message selected at another point.*” This paper established information theory, which had influenced many other disciplines, with applications in data compression, data transmission, speech coding, cryptography, etc. But most of its impact on social sciences only came after Warren Weaver proposed a generalization of the concept communication to encapsulate factors beyond the technical problem, renaming it “*The Mathematical Theory of Communication*” [Shannon et al., 1951]. With Weaver’s interpretation, the communication model now captures the loss of information during the encoding-decoding step. The extended model was widely adopted in areas such as education, communication science, organizational analysis, psychology, etc., which “...*underlies most aspects of our lives*”. Below we briefly describe the Shannon model and the Weaver model.

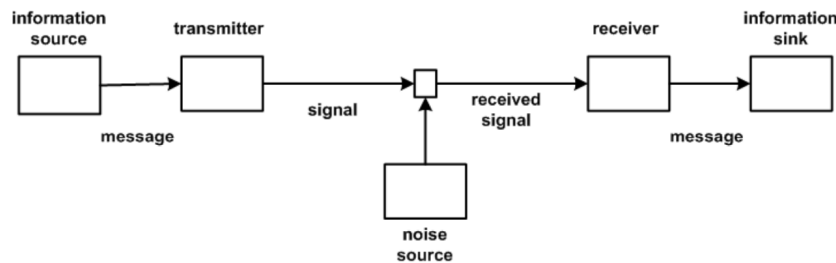


Figure 1.1: Shannon’s Communication Model [Shannon, 1948]

Shannon’s Model. Shannon’s model(Figure 1.1) consists of five components: information source, transmitter, channel, receiver and information sink. The communication starts when an information source is encoded into a message, e.g., it may be encrypted before being transmitted. The encoded message then

travels through a transmitter (channel) and reaches the receiver. The channel can be verbal, written, electronic, audio, video, etc. For example, to tell the same story to an audience, one can use either of a video, a voice message or a picture. As the message travels in the channel, the information transmission may be interfered by noise, e.g., a thunderstorm affects the sound signal. Finally, the message is decoded and passed on to the informee.

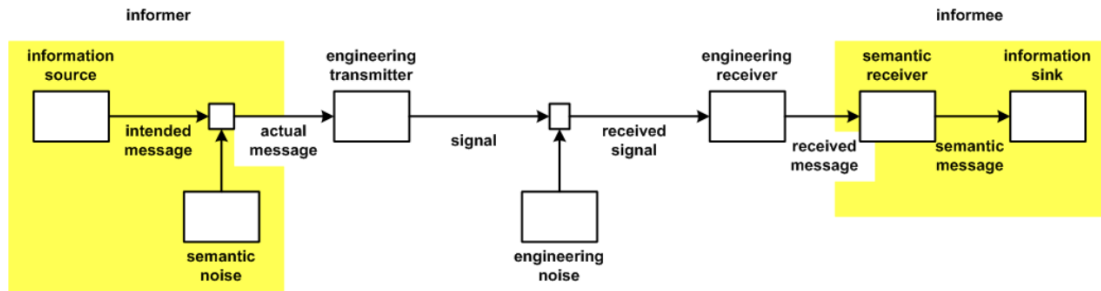


Figure 1.2: Weaver's Communication Model [Shannon et al., 1951]

Weaver's Model. Shannon's model was developed as a mathematical model for the engineering problem of data transmission. Weaver, however, notice that the communication model can be applied to more general scenarios, e.g., communication where one of the informer and informee is a human. When human is involved, more factors which can influence the communication, e.g., human may produce psychological noise during communication [Miller, 1967]. Weaver generalized Shannon's model by asking the following questions:

- A. (*the technical problem*). How accurately can the symbols of communication be transmitted?
- B. (*the semantic problem*). How precisely do the transmitted symbols convey the desired meaning?
- C. (*the effectiveness problem*). How effectively does the received meaning affect conduct in the desired way?

While Shannon's model addressed question A, it does not address question B and C. Weaver's model (Figure 1.2) captured two more factors of noise: first, *semantic noise from informer*. when the message was being encoded, the actual message may be different than the intended message due to semantic reasons (instead of technical reasons), e.g., the message sender does not have the full vocabulary to capture the intended meaning. Second, *semantic noise from informee*. When the message was decoded, the semantic message may again differ from the received message. When the informee is a human, the semantically received message relies on what knowledge the informee possess, how much information the informee could process at the same time, etc. [wea, 2018]. The two noises may occur by itself or at the same time. A simple example is when a non-native speaker talk to a native speaker, vs. when two non-native speakers (from different countries) talk to each other.

A limitation of Shannon-Weaver Model. One limitation of Weaver’s model is that it fails to capture the context in communication, later work extended the model to address more contextual factors [Chandler, 1994]. Essentially, the message, informer, informee and channel may all be the same but have different results of deliveries due to different contexts, such as location. The message *let’s see a movie tonight* may either refer to watching a movie in theater or at home, depending on whether a theater is available in the location.

This thesis focuses on studying the semantic noise during the communication process between a mobile system and the mobile user. That is, one of the informee and informer is the human user, the other is a system that communicates with the user through the channel of mobile user interface. In other words, the theoretical framework this thesis is based on is Weaver’s model, we do not focus on the engineering noise in Shannon’s model.

1.4 System-User Communication in Natural Language

Natural language interface allows users to communicate with the system in an easily understandable way. Although the term natural language interface typically refers to a system that supports natural language input, here we generalize the concept to also include the discussion on natural language output.

1.4.1 Natural Language Input Interface

When used as the input, natural language is mainly for *command* (Section 1.1). The user submits a command in the form of natural language for the system to perform tasks, e.g., web search system, question answering system, speech recognition system, machine translation system, and grammar suggestion systems. Natural language is also used for *satisfy* where the user specifies her requirements in natural language documents.

Under certain scenarios, the contexts in natural language interface is important for the system to understand user intent, so besides the natural language message, the user also passes the context, e.g., demographic information, device, and location.

1.4.2 Natural Language Output Interface

When used as the output, natural language is mainly for *explain*, i.e., providing information about the outside world. The majority of system interfaces are built with natural language, but such interfaces can be optimized for better communication with users using the following techniques:

Summarization: when the original natural language content is too long, the efficiency of communication can be improved by summarizing the content, e.g., search result snippets.

Highlighting: different choices of highlighting words makes a difference in user attention and the effectiveness of browsing natural language interface [Patel et al., 2005].

Adaptation: the design of natural language interface should be adapted to better fit the size of the device, for example, web contents are often reduced to fit mobile screen.

1.5 An Overview of the Mobile Systems and User Interfaces

In this section we give an overview of the mobile systems and user interfaces (channel of communication). Mobile system and mobile devices, in particular smartphon devices, have largely popularized during the last decades. The number of sales in 2018 have been around 1.45 billion, and by the year 2017, more than 31% people worldwide own a smartphone, which had tripled compared with this number in 2011 [sma, 2018]. Such popularity of smartphones is associated with the rise of Android devices, which has dominated the market (77% market share in 2018).

Android system and applications. Android is the open source operating system by Google. The first version of Android was introduced in 2008, the current version is Android 9. Android system is built on top of four layers: the Linux kernel, the library and runtime layer written in C, a Java-compatible application framework, and the application layers. For security, unlike iOS apps which are manually inspected by security vendors, Google has been using the Google Bouncer malware scanner to watch over and scan apps in the Google Play store. Besides, Android security solutions include sandboxing apps, which means that one app does not directly interfere with the operating system or other apps. If an app needs to access certain data (e.g., access to the storage system), it must request the data directly from the user, and the user must directly grant such permission before the app can access the data. The permission system, however, has multiple vulnerabilities, such as stand-alone over-privilege problem and information leakage, inter-process communication (IPC) collusion, permission re-delegation.

Mobile user interfaces. Today's mobile user interfaces all adopt a similar design, which is based on direct manipulation, using touching inputs to correspond to real-world actions, such as swiping, tapping, and with a virtual keyboard. For screen sizes, market statistics show that the sizes of the most popular devices are between 5 to 6 inch, which contributes to 90% of the sales.

1.6 The Challenges of End User Communications Over Mobile Devices

Considering the unique characteristics of mobile devices, including both the mobile system and interface, there are several challenges for a mobile user to communicate with software on the mobile system. Under the Shannon-Weaver model, we are assuming that one side of the informer-informee is the mobile user while the other side is mobile system, and these three characteristics would create/amplify the semantic noises from the user's side:

1.6.1 Challenge 1: Making Security Decisions

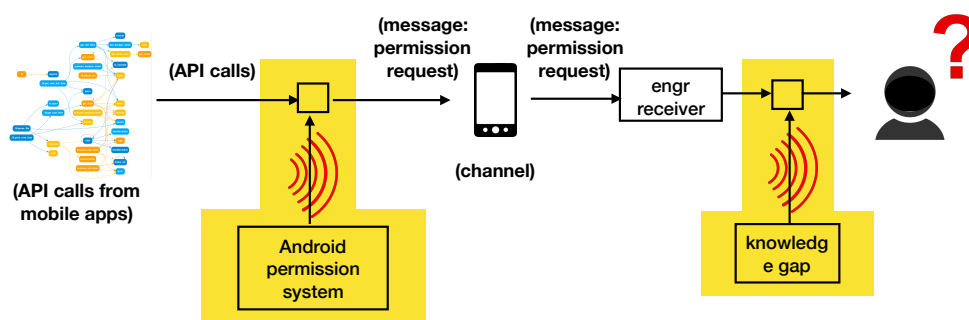


Figure 1.3: Due to user knowledge gap and the simplified permission request, the user gets confused on why the app needs permission request

One reason for the semantic noise is the knowledge gap between the informer (system) and the user (informee). In a mobile system, often the time the user can just use the app without worrying about how things behind the GUI works, e.g., application layer. However, there exist one case where the user needs to make decisions based on her knowledge about the system, i.e., when user needs to make decision on permission requests. Android relies on users to make decisions on the access to her private information. On the one hand, such design is in accordance to the transparency principle, on the other hand, security depends on the context, e.g., when a GPS app requests the location, such request is legitimate while if a gaming app requests the location, such request may be over-privileged. There exists too many apps/functionalities than a system-side access control rule could possibly capture, and as a result, the system relies on the user to make the decision.

To make the permission decision, however, the users need to know why the app needs the permission, and studies show that 1/3 of the time the user cannot understand the reason. In this case, the semantic noise is due to: (1) the message passed contains only the permission request not the purpose of request; (2)

the user does not have the knowledge to decode the purpose behind the permission request.

1.6.2 Challenge 2: e-Commerce Shopping

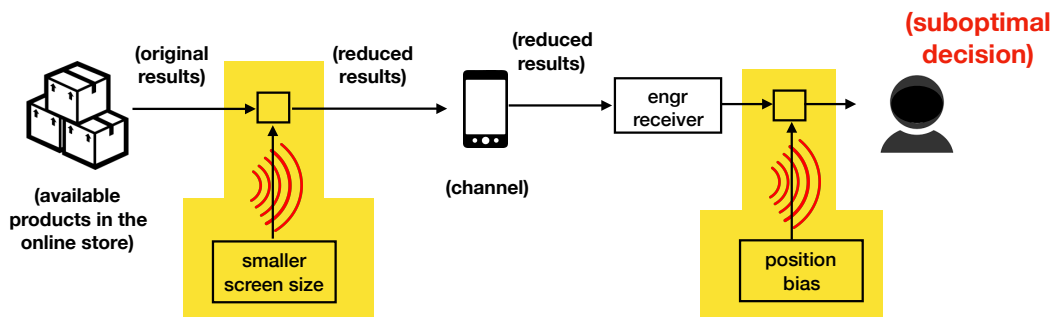


Figure 1.4: Due to reduced search results and user position bias, the user conduct less exploration of the search results, resulting in the user more likely makes a suboptimal decision



Figure 1.5: Difference between display space on laptop/desktop vs. on mobile devices

With the population of mobile devices, more than 40% e-Commerce purchases are made using a mobile phone [mob, 2018b]. However, there exists major challenge for mobile users to explore better options, mostly due to the smaller screen size of mobile devices. As discussed above, most mobile screens are less than 6 inch. The smaller screen gives rise to shrinkage in display space. For example, Figure 1.5 shows how the interface for the shopping website Amazon is different on mobile devices vs. laptop/desktop devices. With the extra space, the laptop/desktop interface can display the item specification including screen size. The structured specification information are effective in drawing user attention compared with the unstructured item title.

In the scenario of online shopping, the informer is Amazon and mobile system and the informee is the user customer. The message passed to the channel are all the retrieved produces from Amazon’s inventory. However, the customer cannot actually receive all the message, because it would be too time and energy consuming to browse all the products. Among the top ranked products, most users would browse less than

30 items, this phenomenon is called users' *search position bias*, meaning they will pay the most attention to the top ranked items, then the attention decays until the search session is eventually abandoned.

Here the position bias is one type of semantic noise due to two facts: (1) the limited space for display on mobile screens; (2) the cognitive limitation of human users to process information. The effect of position bias is that the user fails to gain a comprehensive understanding on the product specifications before making the decision. The user may think the first product is the best one, before realizing there are TVs with larger screen and the same price in the inventory.

Notably, users' position bias also exists on laptop/desktop devices, but the bias on mobile devices is much more severe than that on laptop devices. A study by Google in 2017 [Ong et al., 2017] finds that users browse 50% less products on mobile phones, the input less queries in the same search session, and the query words on mobile phones contain less diversified information. All such results indicate users' search on mobile devices lack *exploration*, which means they tend to quickly adopt a sub-optimal result instead of thoroughly exploring more possibilities before making decision.

1.6.3 Challenge 3: Performing Data Analysis

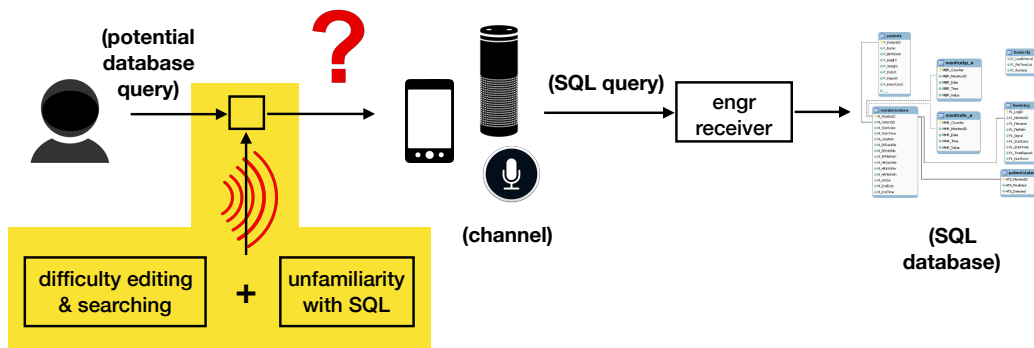


Figure 1.6: Due to the difficulty to write programs using mobile interface, the user's need to query the database is expressed in natural language, which leads to the disconnection to the database interface.

With the population of mobile devices, there is a growing trend of performing data analytics tasks on mobile devices. Companies such as Microsoft, Google, Facebook, Salesforce all launched the mobile app version of their analytics platform, which allows data analytics to perform their tasks on the go.

Due to the touch inputs on mobile devices, however, it is difficult for data analysts to write database queries for analysis. The touch inputs on mobile phone is slower, more error prone (i.e., fat-finger problem [Siek et al., 2005]), also requires more attention than on laptops/desktops [mob, 2018a], This includes the difficulty typing, searching, and copying&pasting, the latter two are essential operations to support

technical inputs. Here the semantic noise comes from: (1) the difficulty for inputting technical information on mobile devices; (2) users' unfamiliarity with database queries (especially beginner analysts).

Summary on the three challenges. All the three challenges happens when user's semantic noise meet the nature of mobile devices: small screens, limitations in input efficiency, and a complex and user-dependent security system.

1.7 Thesis Overview

The goal of this thesis is to assist the communication between end user and mobile systems by realizing the current deficiency factors in the existing systems. Although there might exist other challenges on mobile devices, here we focus on case studying the three challenges discussed in Section 1.6 with the goal of gaining a general understanding of how to improve the communication with mobile devices as the channel. Challenge 2 and Challenge 3 also exists on desktop/laptop, but as mobile devices get more populated today, we find a good motivation to study them under the mobile setting because of the exclusive challenge on mobile devices, a branch of prior work also addressed such exclusive challenge on mobile devices [Popescu et al., 2003, Zhang et al., 2017].

Assisting Security Decision Making. In Chapter 2, we study the challenge for user to make security decisions on mobile devices. First, we analyze the infrastructure of Android security systems on how the permission is requested and why users have difficulty understanding permission requests. Then, we review the existing literatures on user's security understanding, both under mobile and other contexts. Next, to gain understanding on whether mobile apps have provided sufficient education to help user understanding, we conduct an empirical study on such information in existing apps which are created by app developers and security engineers. After such study, we find that the result indicate that there exists a gap that the current supportive information could help for users' understanding. To this end, we finally propose a tool support for assisting app developers and security engineers to better assist users for understanding the security requests, essentially providing more understandable explanations.

Assisting Mobile Browsing and Navigation: In Chapter 3, we study the challenge for users to browse a large database on mobile devices. First, we go through the mobile browsing system, visiting existing solutions for the small screen browsing. Then, we specially discuss the role of *faceted system* for helping browsing and navigation by reviewing solutions in existing literatures. Next, we identify a problem which has not been studied in existing work, which is the numerical facet construction problem. We propose the first system for facet range suggestion. Finally, we identify the challenge in the evaluation of browsing and

propose an interface card model for evaluating the general problem in faceted browsing on mobile devices.

Assisting Database Queries: (to be filled later)

Finally in Chapter 5, we discuss the conclusion that we draw from studying the three cases of challenges.

Chapter 2

Assisting Security Decision Making with Natural Language Explanations

(Below I give an overview of the work I propose to include in this chapter)

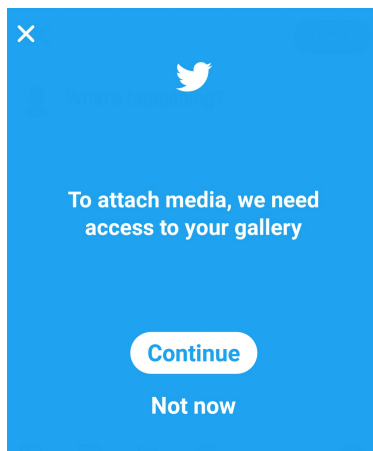


Figure 2.1: Example of Android permission rationale

Studying the Quality of Permission-Explanation Sentences [Liu et al., 2018a]. Because permissions are requested by app developers in the first place, it is not a major challenge for the developers to provide a sentence stating the purpose for the request. The real challenge lies in how to *clearly* explain such permissions to an end user, who is less familiar with Android permissions and APIs. For instance, if the explanation is too abstract, the user may not learn anything new, but if the explanation is too specific, e.g., with many technical details, it can also be difficult for the user to process the technical information. To study the quality of permission-explanation sentences and whether they clearly explain the permission purpose, I conducted a large-scale measurement study on the existing runtime permission rationales in Android apps [Liu et al., 2018a] (an example of such sentences is in Figure 2.1, explaining why the app requested the permission for reading the user’s external storage). I designed a comprehensive methodology for measuring multiple aspects of an explanation’s quality: overall frequency, categorical distributions, correctness, and specificity. For example, Android developers’ guidelines suggest apps to explain permissions that do not look straightforward to users, e.g., a camera app uses users’ location permission (vs. when it uses the camera

permission). To evaluate whether apps are following such guidelines, I examined the statistical correlation between an app’s permission-explanation behavior and the straightforwardness of such permission requests. The results showed that a majority of the studied apps were not following such guideline; in addition, many explanations were mistaken, and a large proportion did not explain any substantial purposes. In summary, the quality of explanations needs to be much improved.

Assisting Apps to Provide Permission Explanations [Liu et al., 2018b]. The low quality of permission explanations in existing Android runtime rationales indicates the challenge for app developers to clearly explain permission purposes to users. To assist developers to improve the quality of their explanations, I built a recommender system named CLAP [Liu et al., 2018b] that recommends candidate sentences for the app developers to refer to. CLAP selects the recommended sentences by conducting information retrieval techniques on the descriptions from the most similar apps. One challenge in this approach was that description sentences from similar apps could not be directly adopted for explaining the current app because they may (partially) describe a different purpose than the current app’s purpose. If many of the recommended sentences were irrelevant, app developers would have to spend a long time manually filtering out these sentences. CLAP used a three-step process to improve the relevance of the recommended sentences: first, CLAP split sentences into smaller units and put all the resulting sentences into a candidate set, so that some sentences in the candidate set may contain less irrelevant information than the original sentence; second, among all the sentences, CLAP ranked candidate sentences by their semantic popularity, i.e., how frequently the sentence’s purpose is mentioned among similar apps; third, CLAP leveraged a few templates to improve the interpretability of recommended sentences. Our experimental results showed that at least 80% of the top-5 recommended sentences matched the ground-truth purpose of the permission request. In addition, our case-study results showed that the top-ranked sentences provided a diverse choices of wording, and these sentences were concise but they explained concrete purposes.

2.1 Overview of Android Security Decision Making

2.1.1 Android Infrastructure and Third-party Market

2.1.2 Android Security System

2.1.3 Over-privileged Problem and User Reaction

2.1.4 Runtime Permission System

2.2 Review Existing Work

2.2.1 Static/dynamic Analysis

2.2.2 Security Risk Assessment

2.2.3 Natural Language Explanation

2.3 An Empirical Study on Natural Language Explanations in Existing Apps

2.3.1 Motivation of Empirical Study

2.3.2 Data Collection Process

2.3.3 RQ1: Population

2.3.4 RQ2: Categorical Explanation Ratio

2.3.5 RQ3: Incorrect Explanations

2.3.6 RQ4: Specificity of Explanations

2.3.7 RQ5: Comparing App Description and Runtime Rationale

2.4 An Analysis of Difficulty of Explaining Permissions

2.5 CLAP: A Recommender System for Assisting Apps Explain Permissions

2.5.1 Module 1: Similar App Retrieval

2.5.2 Module 2: Permission Explanation Detector

2.5.3 Module 3: Sentence Splitter

2.5.4 Module 4: Sentence Reranker

2.5.5 Module 5: Sentence Post-processor

2.6 Experiments on CLAP

2.6.1 Data Collection

2.6.2 Evaluation Metric

2.6.3 Baselines

2.6.4 Experiment 1: Quantitative Study

2.6.5 Experiment 2: Qualitative Study

2.7 Summary

Chapter 3

Assisting Browsing and Navigation in Search Engine

(Below I give an overview of the work I propose to include in this chapter)

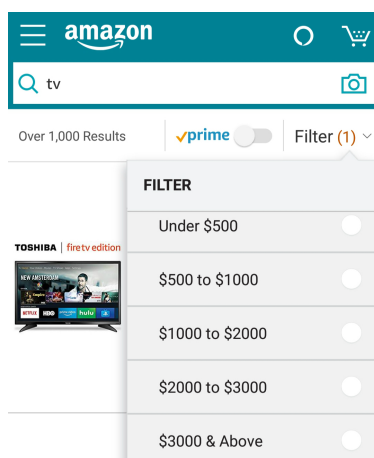


Figure 3.1: Amazon’s numerical range interface on mobile device

Numerical Facets Partition [Liu et al., 2017]. Numerical facets are ubiquitous in end-user application domains, such as price (shopping and travel), distance and rating (local), ages (dating), calories (fitness), and income (jobs). They also widely exist in the domains of software engineering and security, e.g., lines of code (LOC), votes for forum questions, the severity grade of a security vulnerability, and the grade of a student. Compared with other facets (i.e., categorical facets), numerical facets provide a succinct and interpretable summarization for the data distribution. For example, the price ranges in Figure 3.1 tell users that all TVs can be briefly categorized into the five grades, and a large proportion of TVs are under 500\$. In [Liu et al., 2017] we studied the problem of recommending numerical-facet partition to help users with exploratory e-Commerce search. The main challenges in this problem were of two folds: first, how to quantitatively evaluate the performance of the recommended numerical facets; second, how to search for the numerical facets that optimize such evaluation metric. For the first challenge, we proposed to use users’ browsing cost (while using the facets as expansion) as the evaluation metric. With such defined metric, a follow-up challenge was on how to optimize the browsing cost and yet the cost was unknown during testing.

Existing work could be used to estimate the cost, but the accuracy was low. Instead of adopting existing work, we identified that one key step could help solve the problem: a transformation on the math representation of the parameter space. Moreover, some math proofs showed that after this step of transformation, the objective function had an (approximated) upper bound, and the computational time complexity for optimizing the upper bound was largely reduced when compared with that for optimizing the original objective function [Liu et al., 2017]. Furthermore, we developed an advanced model that was more adaptive to different contexts (e.g., query text) based on the regression tree algorithm. Experimental results showed that users can save 25% of efforts when using my system to navigate compared with when using the baseline approach.

Evaluating Mobile Faceted Search [Zhang et al., 2017]. Most of existing work on information retrieval leverages the Cranfield testing methodology, which evaluates a system’s ranking performance on a fixed set of queries and document lists. However, this methodology cannot be used to evaluate today’s search engines, because today users more frequently perform exploratory search, e.g., Google image search now supports exploratory search with tags. To bridge this gap, we proposed a general framework [Zhang et al., 2017] for evaluating the effectiveness of an exploratory search system. To solve the challenges in evaluating a mixture of user operations (e.g., click on document, click on tags, query reformulation) in exploratory search, we jointly modeled these operations using a cost-benefit framework. We tested the integrity of our evaluation framework by comparing its results with the evaluation results from a real user study. Such experiments showed that our framework preserved the consistency of the ordering results between information retrieval systems compared to that from the real user experiment.

3.1 Existing Solutions for the Small-Screen Browsing

3.2 Existing Solutions on Faceted Navigation

3.2.1 Faceted Navigation Systems for Structured Search

3.2.2 Faceted Navigation Systems for Web Search

3.3 Numerical Faceted Search Problem

3.3.1 Applications with Numerical Facet Search

3.3.2 Evaluation Methodology

3.4 Optimizing Numerical Facet Ranges

3.4.1 Baseline 1: Quantile

3.4.2 Baseline 2: Click Model and Dynamic Programming

3.4.3 Parameterizing the Problem

3.4.4 Asymptotic Smoothed Objective Function and An Approximately Upper Bound

3.4.5 Method 1: Context-Ignoring Method

An Overview of Discrete Optimization

Cache-based Algorithm

Analysis of Time Complexity

3.4.6 Method 2: Context-Aware Method

Regression Tree

3.5 Experiments on Numerical Facet Optimization

3.5.1 Dataset Collection

3.5.2 Observation of Asymptotically Smooth

3.5.3 Comparative Results

3.5.4 Ablation Study

3.6 General Evaluation Challenge for Navigational Systems

3.7 A General Framework for Evaluating Navigational Systems

3.8 Summary

Chapter 4

Assisting Database Queries with Natural Language Interface

(this part of work is still ongoing, below I give an overview of the problem definition).

People have started studying the natural language interface for SQL since 1980. The goal is to translate users' natural language queries to SQL queries. This research can help users who is not familiar with SQL language to interact with database, At that time, essentially the growing group of data analysts. There has not been large datasets for training. In recent years, people have created larger datasets, including the recent WikiSQL dataset and the Spider dataset. The Spider dataset is a relatively large dataset which supports many operations including JOIN, GROUP BY, HAVING BY. I plan to study to potentially improve an existing framework, SyntaxSQLNet, on the Spider dataset.

Chapter 5

Conclusion

References

- [goo, 2016] (2016). Google says 20 percent of mobile queries are voice searches. <https://searchengineland.com/google-reveals-20-percent-queries-voice-queries-249917>. Accessed: 2018-12-26.
- [aut, 2018] (2018). Automated individual decision-making, including profiling. <https://gdpr-info.eu/art-22-gdpr/>. Accessed: 2018-12-26.
- [mob, 2018a] (2018a). Continuous input in mobile devices: Pain or gain. <https://www.smashingmagazine.com/2015/03/continuous-input-in-mobile-devices-pain-or-gain/>. Accessed: 2018-12-26.
- [sma, 2018] (2018). Global smartphone shipments forecast from 2010 to 2022 (in million units). <https://www.statista.com/statistics/263441/global-smartphone-shipments-forecast/>. Accessed: 2018-12-26.
- [exp, 2018] (2018). Intelligent machines nvidia lets you peer inside the black box of its self-driving ai. <https://www.technologyreview.com/s/604324/nvidia-lets-you-peer-inside-the-black-box-of-its-self-driving-ai/>. Accessed: 2018-12-26.
- [mob, 2018b] (2018b). Mobile ecommerce stats in 2018 and the future online shopping trends of mcommerce. <https://www.outerboxdesign.com/web-design-articles/mobile-ecommerce-statistics>. Accessed: 2018-12-26.
- [voi, 2018] (2018). Voice search statistics, facts and trends 2019 for online marketers. <https://seoexpertbrad.com/voice-search-statistics/>. Accessed: 2018-12-26.
- [wea, 2018] (2018). Weaver’s model of communication and its implications. <http://www.mrc.uidaho.edu/~rwell/techdocs/Weavers%20Model%20of%20Communication%20and%20Its%20Implications.pdf>. Accessed: 2018-12-26.
- [loc, 2018] (2018). Your apps know where you were last night, and they’re not keeping it secret. <https://www.nytimes.com/interactive/2018/12/10/business/location-data-privacy-apps.html>. Accessed: 2018-12-26.
- [Chandler, 1994] Chandler, D. (1994). The transmission model of communication. *University of Western Australia*. Retrieved, 6:2014.
- [Da Silva and Agusti-Cullell, 2008] Da Silva, F. S. C. and Agusti-Cullell, J. (2008). *Information flow and knowledge sharing*, volume 2. Elsevier.
- [Degeling et al., 2018] Degeling, M., Utz, C., Lentzsch, C., Hosseini, H., Schaub, F., and Holz, T. (2018). We value your privacy... now take some cookies: Measuring the gdpr’s impact on web privacy. *arXiv preprint arXiv:1808.05096*.
- [Hahn et al., 2014] Hahn, L. K., Lippert, L., and Paynton, S. T. (2014). Survey of communication study.

- [Liu et al., 2018a] Liu, X., Leng, Y., Yang, W., Wang, W., Zhai, C., and Xie, T. (2018a). A large-scale empirical study on Android Runtime-Permission Rationale Messages. In *Proceedings of the IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC 2018)*, pages 137–146.
- [Liu et al., 2018b] Liu, X., Leng, Y., Yang, W., Zhai, C., and Xie, T. (2018b). Mining Android app descriptions for permission requirements recommendation. In *Proceedings of the International Requirements Engineering Conference (RE 2018)*, pages 147–158.
- [Liu et al., 2017] Liu, X., Zhai, C., Han, W., and Gungor, O. (2017). Numerical facet range partition: Evaluation metric and methods. In *Proceedings of the International Conference on World Wide Web Companion (WWW 2017)*, pages 662–671.
- [Miller, 1967] Miller, G. A. (1967). The psychology of communication. *Human Resource Management*, 6(3):43.
- [Ong et al., 2017] Ong, K., Järvelin, K., Sanderson, M., and Scholer, F. (2017). Using information scent to understand mobile and desktop web search behavior. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 295–304. ACM.
- [Patel et al., 2005] Patel, A. J. et al. (2005). Systems and methods for highlighting search results. US Patent 6,839,702.
- [Popescu et al., 2003] Popescu, A.-M., Etzioni, O., and Kautz, H. (2003). Towards a theory of natural language interfaces to databases. In *Proceedings of the 8th international conference on Intelligent user interfaces*, pages 149–157. ACM.
- [Shannon, 1948] Shannon, C. E. (1948). A mathematical theory of communication. *Bell system technical journal*, 27(3):379–423.
- [Shannon et al., 1951] Shannon, C. E., Weaver, W., and Burks, A. W. (1951). The mathematical theory of communication.
- [Siek et al., 2005] Siek, K. A., Rogers, Y., and Connelly, K. H. (2005). Fat finger worries: how older and younger users physically interact with pdas. In *IFIP Conference on Human-Computer Interaction*, pages 267–280. Springer.
- [Zhang et al., 2017] Zhang, Y., Liu, X., and Zhai, C. (2017). Information retrieval evaluation as search simulation: A general formal framework for ir evaluation. In *Proceedings of the ACM SIGIR International Conference on Theory of Information Retrieval*, pages 193–200. ACM.