ABSTRACT

Title of Thesis:	ANALYZING USER TRADEOFFS FOR ENCRYPTED EMAIL SERVICES
	Ciara Lynton, Master of Science, 2018
Thesis Directed By:	Professor, Charalampos Papamanthou, Department of Electrical and Computer Engineering

Securing online communication, especially in email settings, is challenging. End-to-end encryption achieves maximal security; however, introducing search capabilities is complicated, potentially making it impractical for email. One option is to locally decrypt and index emails to incorporate search, but this requires significant client-side storage. Encryption that is searchable at the server-side limits local storage, but requires other compromises as well. This thesis presents a study using conjoint analysis to understand user tradeoffs related to email features in order to propose a solution for providing usable, yet secure, email service. The results suggest that while it is ideal to have maximum privacy, users rely heavily on the features present in standard insecure email services. Furthermore, with about half of the participants reporting local device storage as a concern, searchable encryption could be a feasible secure email service solution for some users.

ANALYZING USER TRADEOFFS FOR ENCRYPTED EMAIL SERVICES

by

Ciara Lynton

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Master of Science
2018

Advisory Committee:
Professor Charalampos Papamanthou, Chair
Professor Michelle Mazurek
Professor Tudor Dumitras

Dedication

With God, all things are possible.

To my family, I owe you the world.

Less than one percent of all Electrical Engineering graduate degrees are bestowed upon African-American women [1]. To the less than one percent, our work is not complete. We have a responsibility to grow this number and to encourage our counterparts to do the same.

Acknowledgements

Foremost, I must express my gratitude to my advisor, Professor Charalampos Papamanthou, for his guidance throughout my graduate journey.

I would be remiss if I did not acknowledge Professor Michelle Mazurek for continually going above and beyond with her support during this research project. Moreover, special thanks to Wei Bai for sharing his invaluable insight and for being available day and night to answer all of my questions.

Also, thanks to Professor Tudor Dumitras for serving on my thesis committee.

Lastly, the software used to host this study and to analyze the data was provided by Sawtooth Software, Incorporated. The company granted our research group full access to their state of the art software and for that we are very grateful.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Choice-based conjoint – CBC
Pretty Good Privacy – PGP
End-to-end encryption – E2EE
Amazon Mechanical Turk – MTurk
Human Intelligence Tasks – HITs
Institutional Review Board – IRB
Hierarchical Bayes – HB

Chapter 1: Introduction

1.1 Problem

It is safe to say that society is more connected today than ever before. Whether through the accessibility of reliable transportation, the internet, or countless other communication channels, an individual has the opportunity to connect with just about anyone. These exchanges, however, are not always secure, leaving one vulnerable to various attacks. For instance, online communication channels, such as email, lacking proper security, give email service providers access to personal messages and search queries. This is not a problem until United Healthcare, perhaps, learns of your repeated search query related to a terminal disease and chooses not to provide you with a life insurance policy. In order to prevent such invasions of privacy, developers have created more secure means of online communication that tradeoff between usability and security. This is great for the future of secure communication, but not much is known about how users value the features and whether said tradeoffs would discourage consumers from using the secure services.

1.2 Scope

Focusing on the challenge of securing our email communications and protecting user privacy, this thesis sets out to answer the following motivational questions:

- How can a better understanding of user preferences lead to usable email tools?
- How can researchers determine the best design for secure email service platforms, while maximizing user privacy and considering user preferences?
- What features should developers strongly consider when designing secure email services?
- Would users adapt to a secure email service missing features they consider very important?

We have developed a study using choice-based conjoint (CBC) analysis to determine users' relative level of importance for various email features.

Chapter 2 provides a brief view of related works with Chapter 3 going into greater detail on the usability of currently available email options. Chapter 4 presents the methodology followed by the results and discussion in Chapter 5 and 6 respectively. Lastly, Chapter 7 presents the conclusion and final thoughts on the subject matter.

Chapter 2: Related Work

2.1 Secure Email Options

The public-key encryption scheme was initially designed to secure "electronic mail" [2]. Fast forward 40 years and the standard email service model is still completely insecure. Encryption is not only difficult to implement in email [3], but also in its primitive nature it decreases usability [4]. Steve Sheng et al. evaluated the usability of the PGP 9 (Pretty Good Privacy) email encryption software, finding that even after over 25 years of developing this email encryption tool, there are still significant usability issues [4]. The PGP program has led to the development of various other encrypted email services including the popular Mailvelope end-to-end encryption tool [4]. Orman discusses the portability and usefulness of encrypted email [5].

Gaw et al. preformed an interview study to understand how social factors determine the adoption of encryption technologies [6]. The results suggest that the use of encryption once came with a stigma of paranoia. De Luca et al. preformed a large scale survey and interview study on secure mobile messaging [7]. Their results recognize peer influence as the primary social factor driving people towards secure messaging tools, rather than privacy [7]. Unger et al. identified the privacy, usability

and adoption properties for various secure communication approaches [8]. These works give ground for the argument that security alone is not reason enough for users to embrace secure messaging services.

Song et al. argued the relevance of searchable encryption for mail servers in reducing security risks [9]. They also discussed the provable security of this encryption scheme. Since this initial work, the security and efficiency of searchable encryption has been challenged [10] [11] and improved [12] [13]. Fuller et al. characterized tradeoffs between various searchable encryption solutions [14].

2.3 Typical Email Behaviors

Researchers have devoted significant effort to understanding how users search their emails. Harvey and Elsweiler explored behavioral search patterns in email queries [15]. Whittaker et al. discovered that organizing an inbox does not improve email retrieval success [16]. Cecchinato et al. found that users search their personal and work emails differently [17]. Carmel et al. found that most email search queries in Yahoo! Mail are not suggested by the email service and are typically 1.49 terms long [18]. Likewise Ai et al. reported similar query composition in Outlook [19].

Litmus Software tracked email opens for over 4 years and found that 53% of emails are opened on mobile devices [20]. With the number of individuals owning and

maintaining multiple devices increasing [21], the portability of older messages could obstruct the path to securing email services.

2.3 Conjoint Analysis

Conjoint analysis is a powerful tool for measuring users' relative preferences among features with multiple levels [22]. Using conjoint analysis, Krasnova et al. found that privacy is important to users [23]. This research group used monetary value as the basis of comparison for the various features. Pu et al. used CBC to explore users' online privacy preferences [24]. Burda et al. used CBC questions to determine users' preference towards client-side encryption over server-side encryption [25]. They used monetary value as the basis for their analysis as well.

Turner used the Sawtooth Software to conduct an adaptive conjoint analysis of user preference towards universities [26]. The software also has widespread use across marketing and medical research in determining user preferences [27] [28].

2.4 Preceding Work

The work that this thesis builds upon uses CBC analysis to understand user preferences towards email search features [29]. The results suggest that this method is valid for identifying user preferences towards email features.

This extended work strives to answer the questions left unanswered by the initial work, to validate the findings using the Sawtooth Software and a more comprehensive survey, and to provide a more in depth analysis of the results and implications.

Chapter 3: Usability of Secure Email Approaches

The email service approaches evaluated in this section are the standard email approach (such as Gmail, Yahoo! And Outlook), the classic end-to-end encryption (E2EE) email approach (such as Mailvelope), the local index email approach (similar to the instant messaging services WhatsApp and iMessage), and the cloud index email approach (which includes searchable encryption). These approaches are considered because they are either commercially available or have reasonably efficient performance. Other options, such as oblivious RAM, are not yet practical or scalable for use in email settings [30] [31].

In this section, we will examine the following features as they relate to each secure email service approach:

- **Expressiveness** deals with how much freedom in formulating search queries the email service provides. The expressiveness features we consider in this work are single-word search, multi-word search and partial-word search. Other expressiveness features commonly found in email services include specifying a message's sender, recipient, subject line, date or folder. With the exception of subject line and folder queries in classic E2EE, these features are available in the approaches across the board, but not as commonly used as the aforementioned expressiveness features [18].

- Performance, or simply **storage requirements**, takes into consideration the storage needed on both the client-side and the server-side to support search for the approach. The preceding work considered the costs associated with search based on bandwidth consumption and latency in search and update operations. For standard sized email accounts, all of the options provide efficiency such that the average user would not observe a lag in response time [32] [33].
- **Portability** is described as the ability to access old messages on new devices, or how well the approach scales to more than one device.
- **Security** is defined as how well the email service approach protects users from an adversary interested in learning private message content or search queries. Note that we assume a semi-honest threat model where the service provider is curious, but not actively deviating from the protocol.

Bai et. al analyzed the literature extensively to determine that these are the categories worth examining in this study [29]. The attributes in these categories are important to users in email settings and vary between the email service options.

3.1 Standard Approach

While this option achieves maximal expressiveness, performance, and portability, it does not have acceptable security protocols in place. In fact, users' messages and search queries are collected by their service providers by default. Table 1 displays how standard email fares in each feature category.

Standard		
Expressiveness	Single word search: Multi-word search: Partial word search:	Yes Yes Yes
Storage	Server-side: Client-side:	O(*N*) 0
Portability	Yes	
Security	None	

Table 1: Supported standard email features

3.2 Classic E2EE Approach

Classic E2EE is the opposite of standard privacy in the context of expressiveness. With this approach, messages can only be decrypted locally by the message sender or recipient. In order to achieve this level of privacy, classic E2EE does not support search on the encrypted content of any messages. For this reason, classic E2EE does not require a search index, trivially not consuming any local or cloud storage space for search. This option should not reveal any message content or search queries to the email service provider, so it is considered maximally secure. Table 2 presents the

supported search expressions, performance, portability and security provided by classic E2EE.

Classic E2EE		
Expressiveness	Single word search: Multi-word search: Partial word search:	No No No
Storage	Server-side: Client-side:	N/A N/A
Portability	Potentially	
Security	Maximum	

Table 2: Supported classic E2EE email features

3.3 Cloud Index Approach

Cloud index solutions allow for direct search on encrypted messages. The messages are decrypted locally, then the message keywords are tokenized using a secret key stored on the local device and then mapped to message identifiers. This search index is then securely stored on a server. The cloud storage does not need to be hosted by a separate platform in order to maintain security because the keywords are indistinguishable without the secret tokenization key. To retrieve messages from the email provider, the tokenized version of the keyword is sent to the server managing the search index, and the corresponding message identifiers are returned. The user's device can then retrieve the messages corresponding to the message identifiers from the email provider. This is the way searchable encryption works as proposed by [30] and [33].

Table 3 shows the evaluation of cloud index email service options. Since the search index is stored in the cloud, this option requires local storage just to retrieve the tokenization of each unique word in the messages (O(*W log D*) where W is the total number of unique keywords and D is the total number of messages). The server, however, requires O(*N*) space where N represents the total number of keyword to message pairs. Islam et al. found that the server could learn information about particular search queries using access pattern analysis [31]. Note that cloud index options do not allow for partial word searches. The reason being that partial keywords are not included in the mapping and tokenization of message keywords.

Cloud index		
Expressiveness	Single word search: Multi-word search: Partial word search:	Yes Yes No
Storage	Server-side: Client-side:	O(*N*) O(*W log D*)
Portability	Potentially	
Security	Some	

Table 3: Supported cloud index email features

3.4 Local Index Approach

As with cloud index solutions, local index solutions allow for search on encrypted messages. In this approach, messages are also decrypted locally but the search index is stored in plaintext locally as well. To retrieve messages from the server, this service provides the message identifiers from the locally stored search index table directly. This option does not require any cloud storage for the search table, but it requires local device storage to hold all of the keyword to message mappings. Since all queries are handled locally, no information about the index or queries is given to the email service provider. This ensures added security from the cloud index option in that the search index is taken offline. However, local index solutions cannot provide the same security guarantees as classic E2EE because the email provider can obtain message access patterns. Table 4 presents the features supported by local index.

Local index		
Expressiveness	Single word search: Multi-word search: Partial word search:	Yes Yes Yes
Storage	Server-side: Client-side:	0 O(*N*)
Portability	Potentially	
Security	More	

Table 4: Supported local index email features

Chapter 4: Methodology

4.1 Recruitment

We recruited 253 participants to take our survey on Amazon Mechanical Turk (MTurk). We chose to recruit on MTurk to obtain a large sample of high-quality workers in one convenient space [34]. The primary benefit of using MTurk is in the ability to request workers that meet specific research criteria. To improve the quality of the data, we required our participants to be at least 18 years old, to be located in the United States, to have at least 50 Human Intelligence Tasks (HITs) approved and to have a HIT Approval Rate greater than 95%. The age restriction was to ensure the participants could consent for themselves. We wanted participants living in the United States so that the compensation and currency exchange would be equal. The participants were also limited to taking the survey once. We set the HIT requirements to ensure the participants were in good standing amongst other MTurk recruiters and that they were capable of completing previous surveys [35].

We paid the participants $2.50 to complete the survey, requiring less than 20 minutes of their time. The participants were given a consent form and asked to confirm their age, ability to consent and desire to voluntarily participate in the study.

The study's recruitment procedure was approved by the University of Maryland's Institutional Review Board (IRB).

4.2 Features and Options

We define a feature as a property or an attribute of an email service approach (for example, privacy). We define an option as a setting or a level of a feature (for example, standard, extra or maximum). The features and options were included in the survey based on frequency of use and variance between secure email service options. Single word search, partial word search and multi-word search are the expressiveness features included in the survey. Our previous work did not include single word search expressiveness because it focused primarily on cloud index and local index as viable secure email approaches. To consider classic E2EE, we include single word search as a feature in this study. The search expressiveness difference between cloud and local index approaches is that cloud index does not support partial-word search. Also worth mentioning, additional auxiliary expressions are rarely used in practice [19], so to keep the number of survey questions within reasonable range, we do not include them in the survey.

Classic E2EE and cloud index approaches fit the 5MB device storage option and the local index approach falls into the 500MB device storage option. The actual storage requirements depend on the number of keywords a user has in an email account, so these numbers represent the local storage required for an average-sized email account.

To make the storage requirements relatable, we gave the participants a storage to photo and device application conversion.

Classic E2EE, local index and cloud index approaches have optional portability support. This feature requires a slightly different key management protocol, but it is possible to read messages from before a new device was activated. The cloud and local index approaches would also support search on the older messages.

The privacy feature in the survey specifies what the email service provider can learn about the user's messages and search queries. Standard privacy reflects the standard email approach where the email service provider has access to the contents of all messages and search queries. The extra privacy option echoes the cloud index approach with the email service provider potentially having access to emails and email search queries of high interest. For the definition provided in the survey, both classic E2EE and local index approaches imitate the maximum privacy setting where the email service provider cannot learn the content of any email or search query.

A common attempt to make conjoint utilities more understandable is to express them in dollar equivalents [23] [25]. This is a way of removing the arbitrariness in their scaling [36]. To do this, we include price as a feature. Rather than using a random price for the email service options, we surveyed the costs of email services on multiple mobile device platforms to come up with the free and $1.99 per month prices.

Table 5 presents these features and options and is provided for the participants during the survey.

Features	Feature Descriptions	Options
Price	Monthly fee for using the service	**$0.00** or **$1.99**
Single-word Search	When you search your emails, can you search using a single word? (for example, "hello")	**Yes** or **No**
Multi-word Search	When you search your emails, can you search using multiple words? (for example, "home depot")	**Yes** or **No**
Partial-word Search	When you search your emails, can you search using a partial word instead of a complete word? (for example, "amaz" would return "amazon")	**Yes** or **No**
Privacy	What can your email service provider learn about your emails and email search queries?	**Standard privacy:** Your email provider can read the contents of all email and search queries. This is how most email services currently work (Gmail, Yahoo! Mail, Outlook, etc.). **Extra privacy:** Your email provider can read the contents of certain emails and search queries, but not all. The provider can choose specific topics of high interest to learn about. **Maximum privacy:** Your email provider cannot read the contents of any emails or search queries.
Storage on your device	How much local storage will the service use on your computer or phone?	**5MB on your device:** equivalent to about 2-3 HD photos **500MB on your device:** equivalent to about 200 HD photos or 20 mobile apps
Syncing old email to a new device	After using this email service on one device (a phone or laptop) for several months, you buy a new device. You set up the email service on your new device. Can you read and search old emails on your new device?	**Yes:** read and search all email using the new device both before and after its activation. **No:** read and search only email after you configured the new device. You can't read and search old emails on your new device.

Table 5: Table of features presented in the study

4.3 Conjoint Analysis

As mentioned in Section 2.3, CBC analysis is a powerful tool for determining user's relative preferences towards multiple features.

For our study, we include seven attributes varying between two and three levels. In CBC studies, questions can be presented as either full profiles or partial profiles. Full profile questions display an option from every feature in the study in every profile, while partial profile questions display options for only a subset of the features. Traditional fractional-factorial designs, and specifically the complete enumeration approach, present nearly orthogonal profiles to each participant [37]. We use full profiles and complete enumeration which gives us the option of analyzing each participant's relative preferences individually. With Sawtooth's use of complete enumeration, the full ranking of the 7 features is achieved using just 15 questions with 2 full orthogonal profiles each [38]. This technique provides the equivalent of each participant ranking 192 unique profiles. Prior work has found that survey participants can accurately classify up to 17 profiles before becoming inattentive [39], thus our set of 16 questions fares well.

Questions can either be set for each survey participant or randomized. We use a randomized design to give nearly every participant a unique survey. One of the benefits of the randomized design is that data can be aggregated question-by-question, meaning that for each participant, each response is used to estimate a set of

utilities based on only that question [37]. This allows for high-quality and efficient estimation of aggregate utilities. There will be more information about data analysis in Section 4.5.

The wording of both the questions and the response options in this study is very important. For example, our previous study asked the participants to recommend one of the email service profiles for a friend. This is a standard technique for surveying participants about sensitive topics [40], but it is questionable whether email service options fits in that category. Recommending a service for a friend could disassociate the participant with the profiles, producing nonchalant responses. For this reason, this study requires the participants to choose the email service profile that they would most likely use. Next, the "None" option in a conjoint analysis survey is advised [37], though the hope is that participants do not abused this option. We worded none option as "I am equally inclined to choose either of these options" to ensure the participant realizes that they would have to live with one of the two options, or no email service at all.

Figure 1 shows the prompt presented before the CBC portion of the survey and Figure 2 gives an example of a CBC question from the survey. Note that Table 5 was presented above each CBC question.

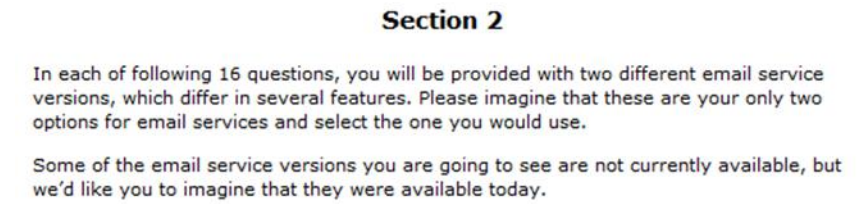

Figure 1: CBC question directions

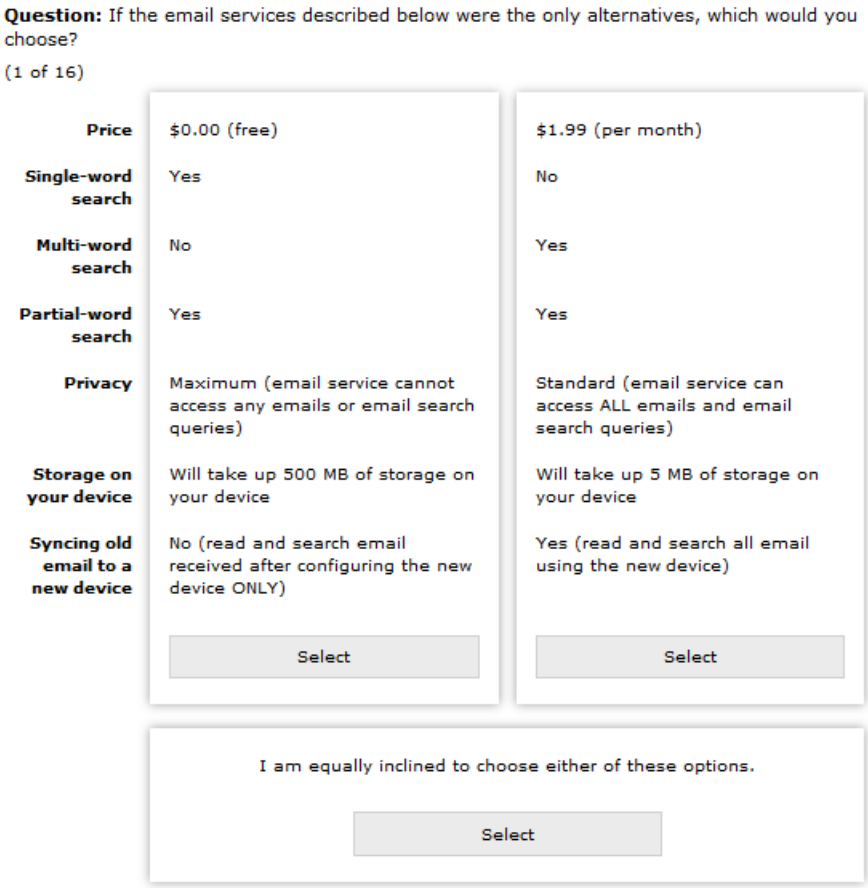

Figure 2: Example of a potential CBC question

While it has been studied that MTurk participants perform well on online attention check questions [41], we included one attention check question during the conjoint analysis portion of the questionnaire. Shown in Figure 3, the fixed CBC question has one profile that is all-around better than the other, encompassing the best option for each feature.

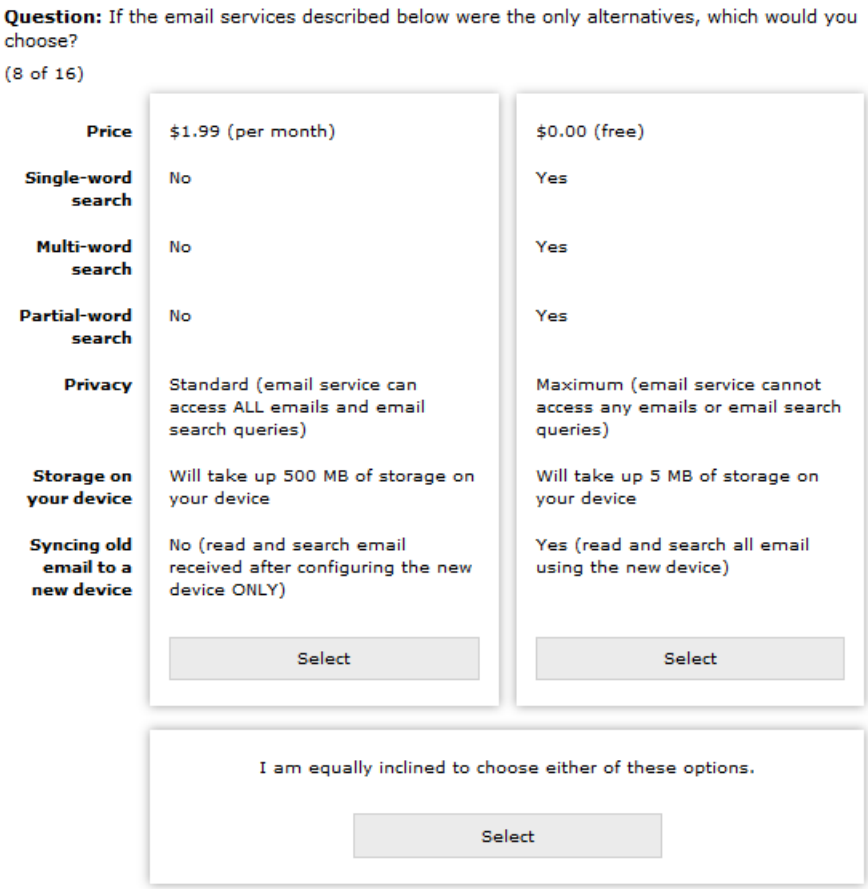

Figure 3: Attention check CBC question

4.4 Clarifying Questions

The survey begins and ends with clarification questions. Figure 4 shows the first question of the survey where participants are given Table 5, the features table, and asked to answer questions based on the feature descriptions in the table. This question is designed to help the participants understand the feature descriptions as presented in the study so they are not making their own assumptions about what the features entail. If the participants answer any of the questions incorrectly, they can review the table and attempt the question again.

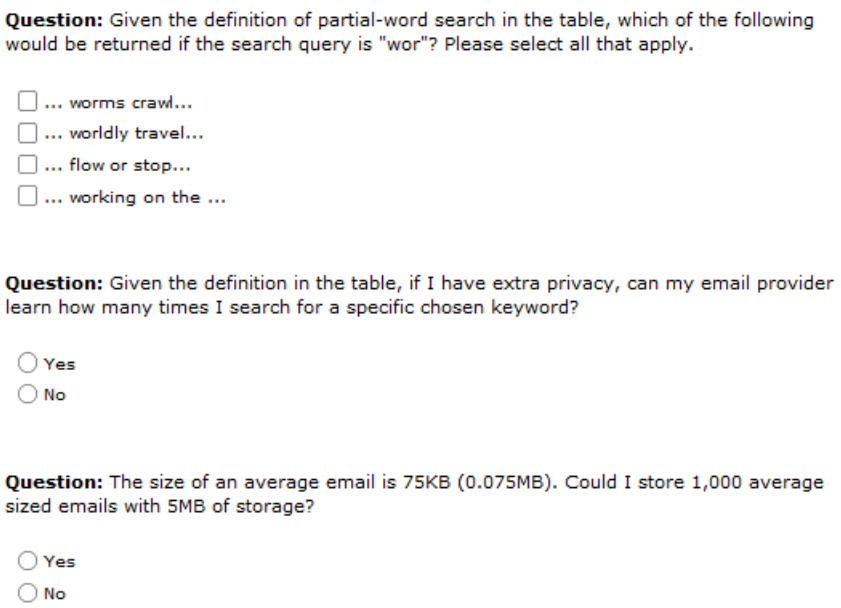

Figure 4: Feature description clarification question

After the CBC questions, participants are asked to report how often they use various features and to specify the features' importance. This question is shown in Figure 5. The responses to this question will allow us to see the correlation between the self-reported importance and the calculated relative preferences.

Section 3

Question: How often do you preform the following tasks?

	Daily	Weekly	Monthly	A few times a year	Once a year or less
Single-word search in your email	○	○	○	○	○
Multi-word search in your email	○	○	○	○	○
Partial-word search in your email	○	○	○	○	○
Access messages on multiple devices	○	○	○	○	○
Delete items on your device to create space	○	○	○	○	○
Search for emails older than 3 months	○	○	○	○	○

Question: When you use the following features, how important is it for you to have them?

	Not at all important	Slightly important	Neither important nor unimportant	Fairly important	Very important
Single-word search	○	○	○	○	○
Multi-word search	○	○	○	○	○
Partial-word search	○	○	○	○	○
Access messages on multiple devices	○	○	○	○	○

Figure 5: Feature use and importance clarification question

An additional question, shown in Figure 6, was only presented to 100 random participants. The question asks if the participants would use a service missing various features, but including maximum privacy. This aims to answer whether users would adapt to any secure email service they are given, for the sake of privacy. Based on the responses to these clarification questions, participants are given 1 open-ended question asking why they selected a feature as very important.

Section 3

Question: In order to obtain maximum privacy, how likely are you to use an email service that **DOES NOT HAVE** the following features?

	Very likely (would use the service)	Likely	Indifferent	Unlikely	Very unlikely (would not use the service)
Single-word search	○	○	○	○	○
Multi-word search	○	○	○	○	○
Partial-word search	○	○	○	○	○
Access messages on multiple devices	○	○	○	○	○

Figure 6: Missing feature clarification question

Next, participants report how well they understand a set of six web topics based off of the well-validated Hargittai et al. scale [42] by answering the question shown in Figure 7. Lastly, we collect demographic information from each participant and this, along with the web skills question, is used to determine if there are any direct correlations between technical or socioeconomic backgrounds and the participants' email feature preferences.

Section 4

Question: There are 6 items in this section. How familiar are you with the following computer and Internet-related items? Please choose a number between 1 and 5, where 1 represents no understanding and 5 represents full understanding of the item.

	1 (No Understanding)	2	3	4	5 (Full Understanding)
Advanced search	○	○	○	○	○
PDF	○	○	○	○	○
Spyware	○	○	○	○	○
Wiki	○	○	○	○	○
Cache	○	○	○	○	○
Phishing	○	○	○	○	○

Figure 7: Web skills clarification question

The responses to the questions in this section are used to understand the "why" behind participants' relative feature preferences.

4.5 Data Analysis

There are a number of ways to analyze CBC data. We could use the counting model, simply considering the number of times an option level was chosen versus the number of times the option level was presented. There is also the logit method which uses the participant's part-worths, or conjoint utilities, to determine the probabilities of the participant choosing alternative options. Then there is the Hierarchical Bayes (HB) model which leverages information from all participants to estimate the results for each individual participant [43]. This model is thorough in that it collects each participant's part-worths in a multivariate normal distribution using Gibbs Sampling [43]. This distribution is characterized by a vector of means and a matrix of covariances [43]. For our analysis, we will use both the count and HB models.

Since we want to characterize the relative importance of each feature, we will consider the difference each feature makes in the total utility of an email service. Moreover, we also want to characterize the relative importance of each feature option, so we will consider the difference each feature option makes in the total utility of each profile. We will then calculate the percentages from the relative ranges to obtain a set of feature importance totaling to 100 percent.

We will use various statistical analysis techniques to interpret the data as well. Our predictive model used to obtain the correlation between the features is regression. To identify correlations between chosen options and self-reported data we use segmentation by way of convergent cluster and ensemble Analysis. This is a way of categorizing participants into groups in order to determine if the survey respondents "cluster" into identifiable groups [44].

4.6 Limitations

As with any study, there are limitations to our methodology. First, the population of people who use email is not completely reflected by the distribution of survey participants. In fact, MTurk workers are typically younger, predominantly white, more educated and more concerned about online privacy than the general population [45]. Furthermore, the primary concern lies in the question of whether self-reported data is enough to draw any real conclusions. Vetschera and Kainz found that even with anonymous surveys, participants sometimes inflate the characteristics they feel seem "right" [46]. This is somewhat stabilized in our study with the use of CBC analysis. Participants are not explicitly choosing between one feature and another, but rather they are choosing between profiles. Therefore, a participant would have to go to greater lengths to disguise their true preferences. This study may not be the basis of all future secure email service decisions, but it will give the developers insight to user preferences and attitudes towards email service features.

Chapter 5: Results

This chapter presents the findings of our study. These findings are then analyzed in Chapter 6.

5.1 Participants

5.1.1 Demographics

Of the 253 participants recruited on MTurk, we received 200 uniquely qualified surveys. Identifying careless and inattentive survey data is critical to establishing accurate results [47]. Table 6 shows a breakdown of the disqualified questionnaire respondents. To select the cutoffs for disqualification criteria we preformed time experiments on 12 volunteers with diverse backgrounds. We first had six volunteers attentively take the survey and they performed this task in an average of 15 minutes, ranging between 12 minutes and 17 minutes. Next, we had six volunteers attempt to take the survey as quickly and as accurately as possible. These participants completed the survey in times ranging between eight minutes and 14 minutes. They completed the conjoint analysis section in times ranging between 2 minutes and 5 minutes. Since MTurkers are considered more tech-savvy and spend more time completing online surveys than our volunteer pool, we took the minimum times demonstrated by this sample and reduced it by 25% to come up with the disqualification bounds.

Hence, participants who completed the survey in less than six minutes, completed the conjoint analysis portion of the survey in less than 100 seconds, answered any of the conjoint analysis questions in under three seconds, answered the attention check question inaccurately or took the survey multiple times were disqualified. The table also shows the overlap of participants who were disqualified in multiple categories. 50% of disqualified participants were ineligible in more than one category, further validating their removal and the disqualification criteria.

	Participant Number	**Total number of participants**	**Number of participants delinquent in other categories**
Survey completed in less than 6 minutes	30, 31, 35, 48, 50, 61, 106, 121, 134, 164, 166, 178, 188, 189, 194, 198, 200, 212, 223, 227, 236, 252, 257, 275, 277, 284, 293, 301, 304	29	21 = 72.4%
Conjoint analysis section completed in less than 100 seconds	35, 106, 164, 166, 198, 203, 227, 252, 277, 282, 293	11	11 = 100%
Conjoint analysis profile selected in 3 seconds or less	26, 30, 31, 35, 47, 55, 61, 106, 112, 113, 121, 128, 150, 164, 166, 188, 198, 200, 202, 203, 209, 218, 227, 236, 252, 254, 257, 263, 268, 271, 273, 275, 282, 283, 284, 293, 301	37	28 = 75.8%
Wrong attention check choice selected	48, 92, 135, 136, 166, 175, 183, 184, 188, 200, 203, 236, 252, 254, 268, 282	16	10 = 62.5%
Completed the survey multiple times	55 = 164 64 = 112 and 113 200 = 202 and 203 268 = 271 and 275	11	10 = 90.9%

Table 6: Breakdown of the disqualified questionnaire responses showing the overlap in disqualified participants based on more than one criteria

Table 7 shows the demographics of the eligible participants. 56% of the participants were male, the mean age reported was 37 years old, 81% of the participants were white, 37% have completed at least their Bachelor's degree, 46% have an income above 50,000 dollars, 21% work or were educated in a technical field and of the 21% of participants with technical backgrounds, 7% specialize in security.

Gender	Male	111	56%
	Female	87	44%
	Other	1	<1%
	Prefer not to answer	1	<1%
Age	18-29	44	18%
	30-39	89	35%
	40-49	39	16%
	50-59	18	7%
	60-69	8	3%
	70-79	2	1%
	80+	0	0%
Race	American Indian or Alaska Native	1	<1%
	Asian	11	6%
	Black or African American	19	10%
	Hispanic or Latino	13	7%
	Native Hawaiian or Other Pacific Islander	2	<1%
	White	152	76%
	Other	1	<1%
	Prefer not to answer	1	<1%
Education	Some high school	1	1%
	High school or GED	16	8%
	Some college	59	30%
	Trade/technical/vocational training	4	2%
	Associate's Degree	31	15%
	Bachelor's Degree	74	37%
	Master's Degree	12	6%
	Professional degree	3	1%
	Doctorate degree	0	0%
Technical background	Yes	41	20%
	No	150	75%
	Prefer not to answer	9	5%
Security background (If technical)	Yes	3	7%
	No	38	93%
Income	Less than $25,000	45	23%
	$25,000 to $49,999	60	30%
	$50,000 to $74,999	37	19%
	$75,000 to $99,999	29	14%
	$100,000 to $149,999	20	10%
	$150,000 or more	4	2%
	Prefer not to answer	5	2%

Table 7: Participant demographics (the percentages may not add up to 100% because participants had the option of not answering these questions)

5.1.2 Web Skills

Table 8 displays the self-reported web skills of the 200 participants. 44% report having full understanding of advanced search (mean=4.02), 55% for PDF (mean=4.39), 39% for spyware (mean=4.00), 51% for Wiki (mean=4.17), 39% for cache (mean=3.96) and 43% for phishing (mean=3.96). On the 5-point scale, these participants have a score of 4.1.

Advanced search	1 (No understanding)	2	1%
	2	17	9%
	3	44	22%
	4	50	25%
	5 (Full understanding)	87	44%
PDF	1 (No understanding)	2	1%
	2	4	2%
	3	19	10%
	4	65	33%
	5 (Full understanding)	110	55%
Spyware	1 (No understanding)	4	2%
	2	15	8%
	3	37	19%
	4	66	33%
	5 (Full understanding)	78	39%
Wiki	1 (No understanding)	9	5%
	2	6	3%
	3	28	14%
	4	56	28%
	5 (Full understanding)	101	51%
Cache	1 (No understanding)	4	2%
	2	22	11%
	3	29	15%
	4	68	34%
	5 (Full understanding)	77	39%
Phishing	1 (No understanding)	6	3%
	2	19	10%
	3	38	19%
	4	52	26%
	5 (Full understanding)	84	43%

Table 8: Participant web skills

5.2 Conjoint Analysis

Figure 8 presents a pie chart of the HB analysis data for the 200 participants. The relative importance ranks price as most important with 20.2%, followed by privacy with 18.7%, message portability with 13.8%, storage with 13.5%, multi-word search with 12.5%, single-word search with 11.6% and lastly partial-word search with 9.7%.

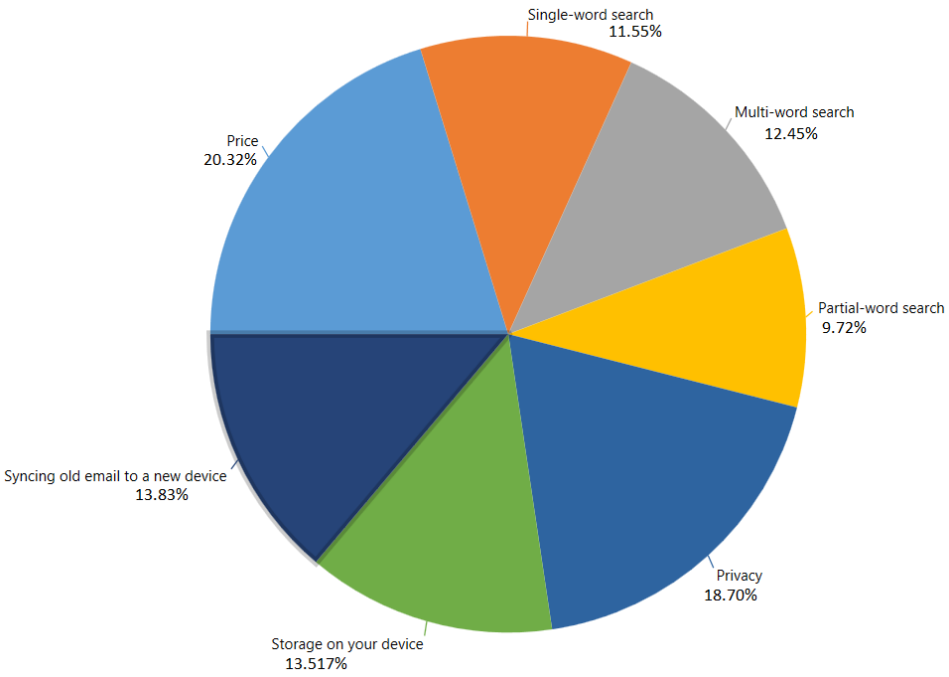

Figure 8: Pie chart for the relative importance of each email feature

Table 9 presents the part-worth utilities and relative importance for the features and Figure 9 shows the cumulative distribution of the relative importance for each participant.

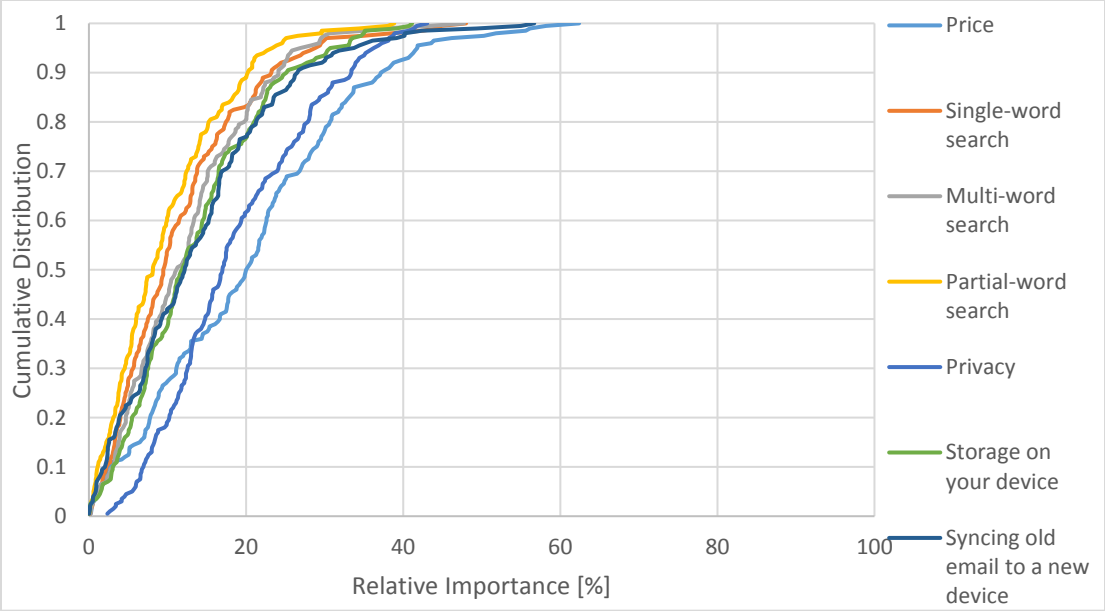

Figure 9: Cumulative distribution of the relative importance of each feature for each participant

Features	Options	Part-worth utility	Lower 95% CI	Upper 95% CI	Relative importance	Std deviation	Count Ratio
Price	$0.00 (free)	0.253	0.205	0.301	20.232%	0.347	0.650
	$1.99 per month	-0.253	-0.301	-0.205		0.347	0.293
Single-word search	Yes	0.084	0.053	0.115	11.550%	0.221	0.535
	No	-0.084	-0.115	-0.053		0.221	0.408
Multi-word search	Yes	0.074	0.041	0.107	12.448%	0.240	0.521
	No	-0.074	-0.107	-0.041		0.240	0.422
Partial-word search	Yes	0.034	0.005	0.062	9.720%	0.207	0.523
	No	-0.034	-0.062	-0.005		0.207	0.420
Privacy	Standard privacy	-0.091	-0.132	-0.051	18.699%	0.295	0.351
	Extra privacy	-0.002	-0.034	0.031		0.236	0.477
	Maximum privacy	0.093	0.053	0.134		0.292	0.588
Storage	5 MB	0.030	-0.066	0.005	13.517%	0.255	0.486
	500 MB	-0.030	-0.005	0.066		0.255	0.458
Portability	Yes	0.055	0.016	0.094	13.834%	0.283	0.527
	No	-0.055	-0.094	-0.016		0.283	0.416

Table 9: Part-worth utilities, upper and lower confidence intervals, relative importance, standard deviations and count ratios

Table 10 displays the part-worth utility change for the feature options and the dollar equivalents. To find the dollar equivalents we divided the utility change of the price feature (0.506) by the price ($1.99) and then divided each utility change by that number (0.25427).

Features	Option Change	Utility Change	Dollar Equivalent
Price	$1.99 per month →$0.00 (free)	0.506	
Single-word search	No → Yes	0.168	$0.66
Multi-word search	No → Yes	0.148	$0.58
Partial-word search	No → Yes	.067	$0.26
Privacy	Standard→Maximum	0.184	$0.72
	Extra → Maximum	0.095	$0.37
	Standard → Extra	0.089	$0.35
Storage	500 MB → 5MB	0.060	$0.24
Portability	No → Yes	0.110	$0.43

Table 10: Part-worth utility change for the features and options and the dollar equivalent of the changes

5.3 Email Feature Use and Importance

5.3.1 Feature Use

Tables 11, 12 and 13 present the data for how often the participants use single-word search, multi-word search and partial-word search respectively. On at least a monthly basis, 79% of the participants use single-word search, 61% use multi-word search and 53% use partial-word search.

Frequency of use: single-word search			
Value	Label	Count	Percent
1	Daily	22	11
2	Weekly	85	43
3	Monthly	49	25
4	A few times a year	28	13
5	Once a year or less	16	8

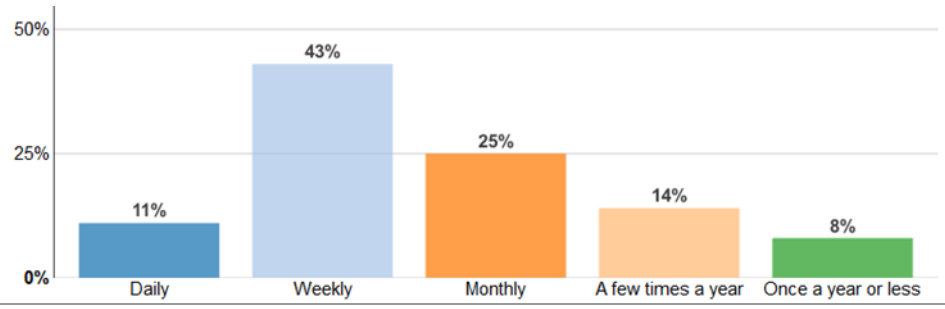

Table 11: Frequency of use for the single-word search feature

Frequency of use: multi-word search			
Value	Label	Count	Percent
1	Daily	16	8
2	Weekly	59	30
3	Monthly	45	23
4	A few times a year	37	18
5	Once a year or less	43	21

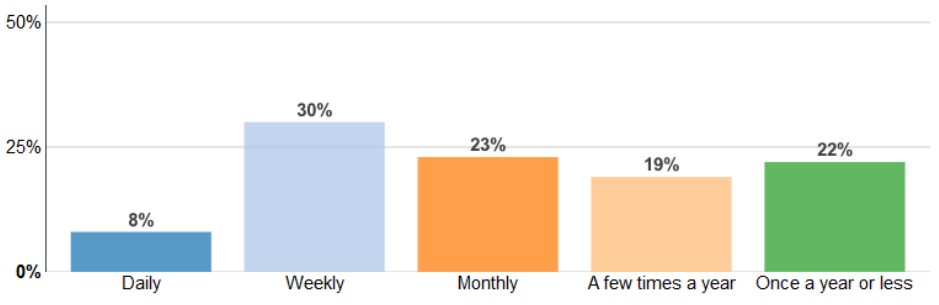

Table 12: Frequency of use for the multi-word search feature

Frequency of use: partial-word search			
Value	Label	Count	Percent
1	Daily	14	7
2	Weekly	48	24
3	Monthly	43	22
4	A few times a year	46	23
5	Once a year or less	49	24

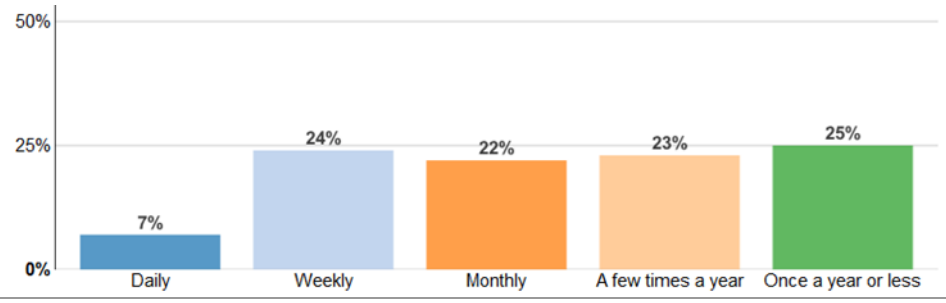

Table 13: Frequency of use for the partial-word search feature

Table 14 shows how often the participants access emails on more than one device, with 90% of participants performing this task at least monthly.

Frequency of use: access emails on multiple devices			
Value	Label	Count	Percent
1	Daily	99	50
2	Weekly	55	28
3	Monthly	24	12
4	A few times a year	11	5
5	Once a year or less	11	5

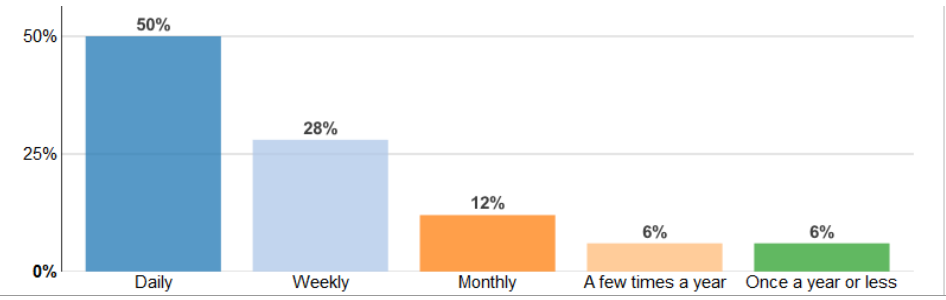

Table 14: Frequency of access to emails on multiple devices

Table 15 shows how often the participants delete items on their device to save space, with 53% of the participants performing this task at least monthly. Furthermore, 45% of survey participants report local device storage as a problem for their devices.

Frequency of use: delete items on device to save space			
Value	Label	Count	Percent
1	Daily	20	10
2	Weekly	33	17
3	Monthly	52	26
4	A few times a year	43	21
5	Once a year or less	52	26

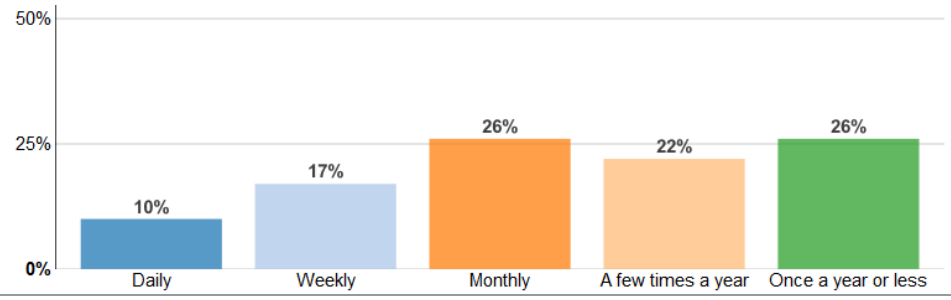

Table 15: Frequency of deleting items on a device to save space

Table 16 illustrates how often the participants search for emails older than 3 months. 41% of the participants perform this task at least monthly.

Frequency of use: search for emails older than 3 months			
Value	Label	Count	Percent
1	Daily	3	3
2	Weekly	22	10
3	Monthly	56	28
4	A few times a year	97	49
5	Once a year or less	22	10

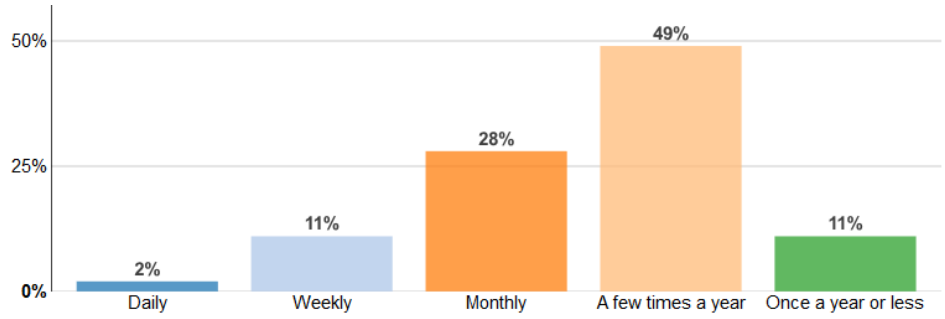

Table 16: Frequency of searching for emails older than 3 months

5.3.2 Feature Importance

Tables 17, 18 and 19 present the distribution for how important participants rate single-word search, multi-word search and partial-word search respectively. 63% of the participants find single-word search important, 45% find multi-word search important, and 38% find partial-word search important.

Feature importance: single-word search			
Value	Label	Count	Percent
1	Not at all important	18	8
2	Slightly important	45	23
3	Neither important nor unimportant	12	6
4	Fairly important	75	38
5	Very important	50	25

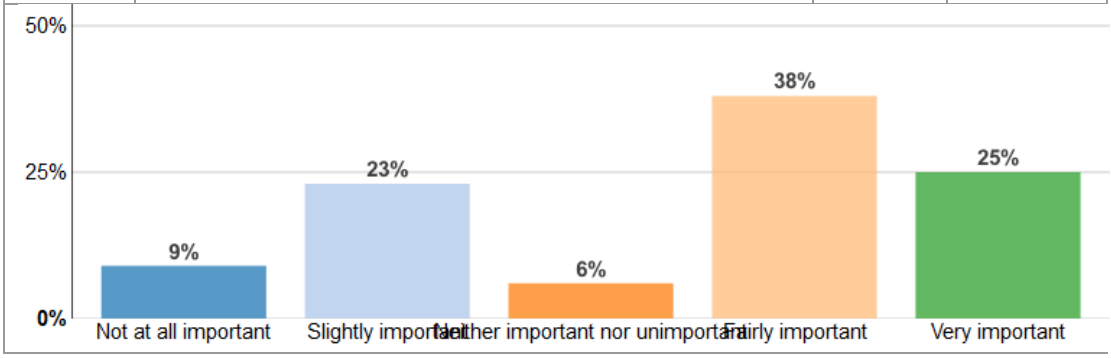

Table 17: Importance of single-word search

Feature importance: multi-word search			
Value	Label	Count	Percent
1	Not at all important	33	16
2	Slightly important	50	25
3	Neither important nor unimportant	29	14
4	Fairly important	57	29
5	Very important	31	16

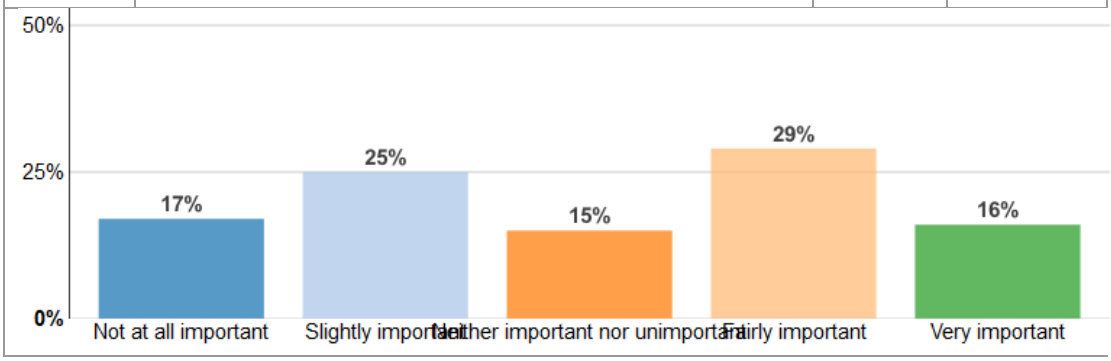

Table 18: Importance of multi-word search

Feature importance: partial-word search			
Value	Label	Count	Percent
1	Not at all important	47	23
2	Slightly important	48	24
3	Neither important nor unimportant	30	15
4	Fairly important	54	27
5	Very important	21	11

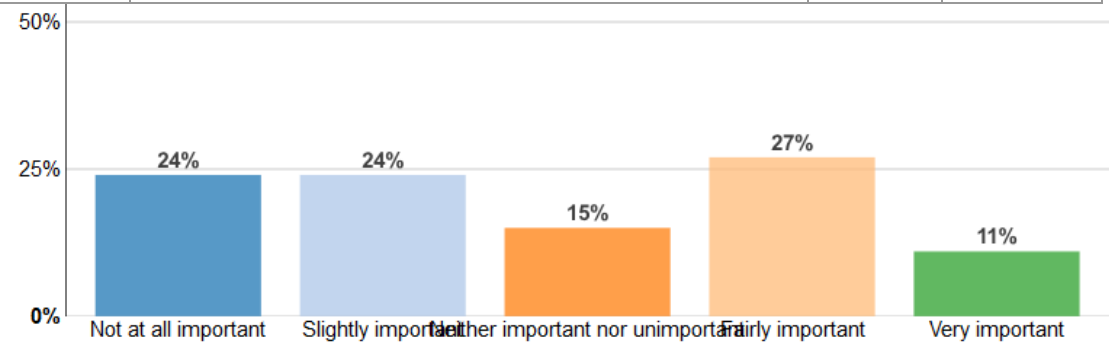

Table 19: Importance of partial-word search

Table 20 illustrates the distribution for how important participants rank having access to messages on more than one device. 74% of the participants find this portability feature important.

Feature importance: access to emails on multiple devices			
Value	Label	Count	Percent
1	Not at all important	8	4
2	Slightly important	22	11
3	Neither important nor unimportant	23	12
4	Fairly important	52	26
5	Very important	95	48

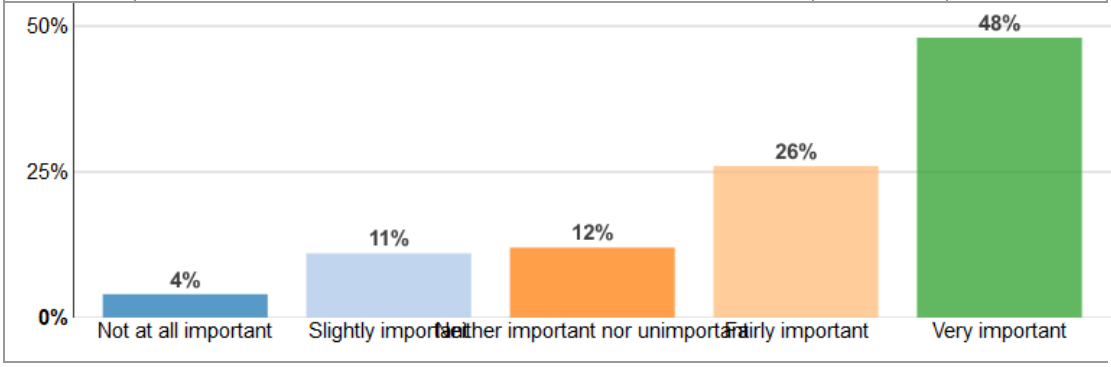

Table 20: Importance of access to emails on multiple devices

5.3.3 Missing Features

Select participants were asked how likely they would be to use an email service that has maximum privacy, but is missing either single-word search, multi-word search, partial-word search or message portability. Tables 21, 22, 23 and 24 present the data for the 73 eligible participants. 51% of the participants are unlikely to use an email service without single-word search, 27% for an email service missing multi-word search, 21% for an email service lacking partial-word search and 53% for an email service without access to emails on multiple devices.

Missing features w/ maximum privacy: single-word search			
Value	Label	Count	Percent
1	Very likely (would use the service)	11	15
2	Likely	11	15
3	Indifferent	14	19
4	Unlikely	28	39
5	Very unlikely (would not use the service)	9	12

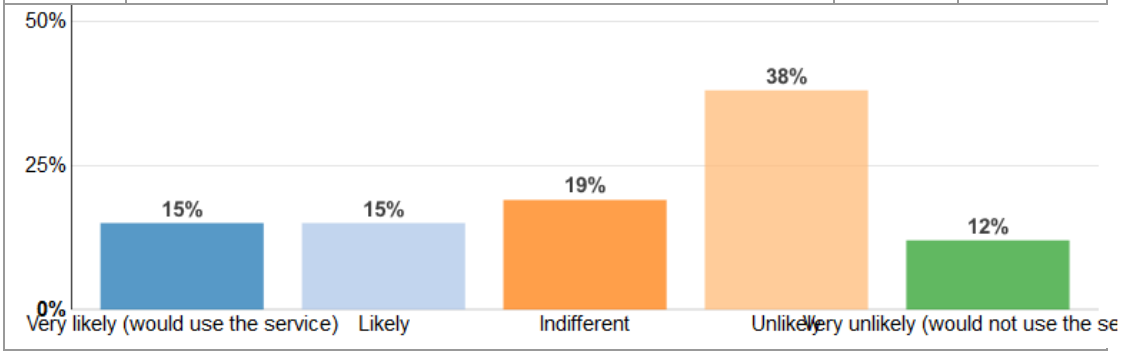

Table 21: Likelihood of using an email service with maximum privacy, but without single-word search

Missing features w/ maximum privacy: multi-word search			
Value	Label	Count	Percent
1	Very likely (would use the service)	16	22
2	Likely	16	22
3	Indifferent	21	29
4	Unlikely	16	22
5	Very unlikely (would not use the service)	4	5

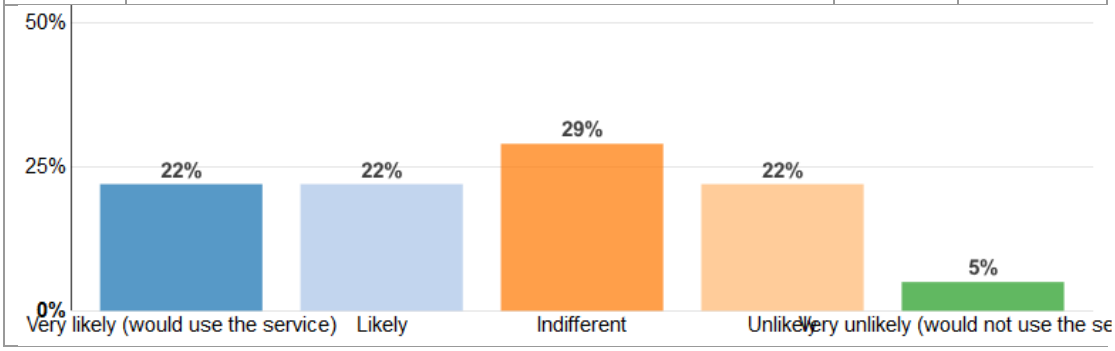

Table 22: Likelihood of using an email service with maximum privacy, but without multi-word search

Missing features w/ maximum privacy: partial-word search			
Value	Label	Count	Percent
1	Very likely (would use the service)	16	22
2	Likely	17	23
3	Indifferent	25	34
4	Unlikely	11	15
5	Very unlikely (would not use the service)	4	6

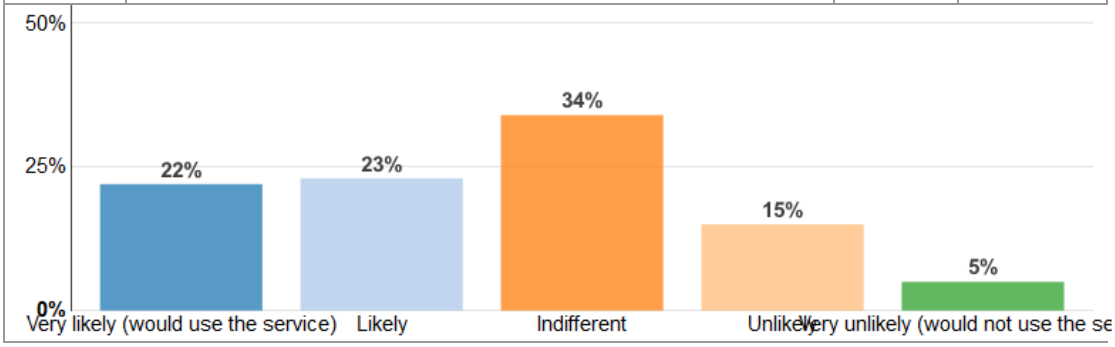

Table 23: Likelihood of using an email service with maximum privacy, but without partial-word search

Missing features w/ maximum privacy: access to email on multiple devices			
Value	Label	Count	Percent
1	Very likely (would use the service)	7	10
2	Likely	16	22
3	Indifferent	11	15
4	Unlikely	22	30
5	Very unlikely (would not use the service)	17	23

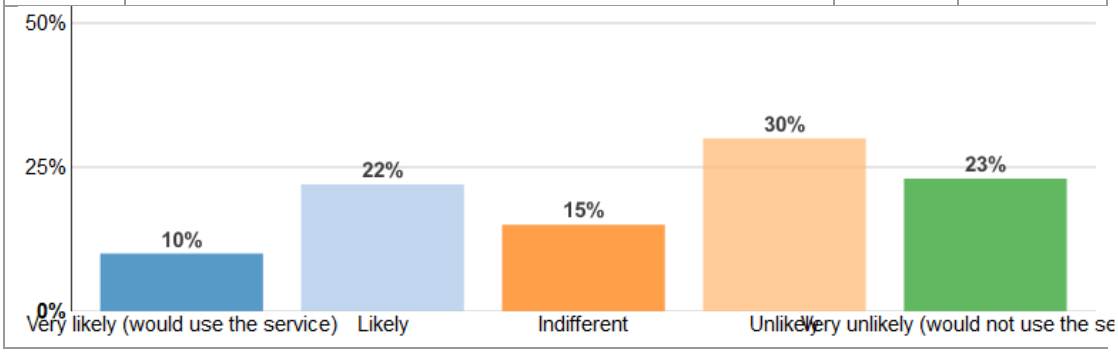

Table 24: Likelihood of using an email service with maximum privacy, but without access to emails on multiple devices

Chapter 6: Discussion

In this chapter we review the results found in Chapter 5 and make connections between various entities. Unless otherwise noted, the information presented in this chapter has been determined statistically significant with $p < 0.01$ using Sawtooth's count analysis tool.

We use multiple statistical analysis techniques to interpret the data. The predictive model used to get the correlations between the features is regression. This tool takes as input all of the user responses to the CBC questions and normalizes them to give the relative importance for each feature. To identify correlations between chosen options and self-reported data we use segmentation by way of Convergent Cluster and Ensemble Analysis. This method takes as input the utilities, importance, demographics, web skills, feature use or feature importance. It then categorizes respondents into sets based on these inputs in order to determine if the participants "cluster" into identifiable and statistically important groups [44]. These tools are built into the Sawtooth Software.

6.1 Participants

The participants who took this survey fit the MTurk pattern of being more tech-savvy. Though only 21% report having technical backgrounds, they had an average score of 4.1 on the web skills scale which is 14% higher than the score of 3.4 reported by Hargittai et al. [42]. This could also be attributed to the fact that this study was performed 8 years ago. It is interesting to note that 91% of the participants at least started college.

6.2 Conjoint Analysis

Overall, the CBC HB results make sense and align with the results of the previous study.

As with that study, price and privacy have the highest relative importance. The count ratio is the ratio between the number of times an option is shown and the number of times the option is selected. The count ratio of the $0.00 (free) option is 0.650 and for the $1.99 per month option it is 0.293. This means that when the paid option was presented, participants only selected the profile with this option 29% of the time. This feature has the largest gap between options with a 35.7% difference between the free and paid options. We can infer that some users are not willing to pay for an email service. This is likely attributed to the fact that there are a plethora of free email services currently available.

The count ratio for standard privacy is 0.351, for extra privacy it is 0.477 and for maximum privacy it is 0.588. Moreover, the utility change for standard privacy to maximum privacy is 0.184 ($0.72), for extra privacy to maximum privacy it is 0.095 ($0.37) and for standard privacy to extra privacy it is 0.089 ($0.35). Relative to the standard to maximum privacy utility change of 0.184, the extra to maximum privacy difference is 48% and the standard to extra privacy difference is 52%. This 4% difference implies that jumping from standard to extra privacy is only slightly less valued than the jump from extra privacy to maximum privacy ($0.02). In our previous study, standard to extra privacy had a greater utility change than extra to maximum privacy, though neither of the utility changes came close to the utility change for standard to maximum privacy. This implies that the participants in this study value maximum privacy more than the participants in our last study though altogether, these results verify the results found in our previous study that extra privacy may be good enough for most users.

Not only is the relative importance for price and privacy highest, but the standard deviation for these features is also the highest. This means that users have a higher variance in their views of importance for these features then, say, partial-word search which has the lowest variance. It seems that users can come to a consensus that certain features are not important, but the features that are most important vary.

Portability and storage have almost equal relative importance, 13.8% and 13.5% respectively. The count ratio for 5MB of storage and 500MB of storage is almost

equal, with 5MB having only a 0.028 higher ratio, and both are less than half (statistically insignificant). A possible explanation for the ratios being so close is that both 5MB of storage and 500MB of storage are acceptable to some of the participants. This is verified by the distribution of the relative importance of storage for each participant. The range for the relative importance of storage is from 0.01 to 41.14 with the median being 11.84. That means that half of the participants had storage relative importance less than 11.84 and half had it more than 11.84. Since the results are aggregated, the total relative importance reflects an average. Therefore, we see that some participants heavily considered storage when choosing between profiles and others did not consider it as heavily. This explains why the standard deviation and relative importance is high, but the part-worth is low. Moreover, a theory for why both options have ratios less than one half is that storage is the factor most causally related to participants choosing the "None" option. We will discuss the "None" statistics later in this section.

Single-word search and multi-word search have close relative importance, differing by less than 1%. The count ratio for these search features is about one-half. This could be because there were likely a number of times a profile had one of the search features, but not the other. Meaning, it is easier to live without single-word search if an email service provides multi-word search.

Besides price, privacy and partial-word search, the other features had comparatively close relative importance, each varying by less than 2.5%. This could imply that these

features are all important to users, and therefore important to include in a secure email service. The low relative importance of partial-word search may mean that single-word search and multi-word search are sufficient for searching emails.

The "None" option was indeed use sparingly in this study. 75 of the 200 qualified participants used this option at least once. Of the 75 participants who used this option, the mean number of times it was used in a participant's CBC data set is 2.3 or 14.4% of the time (statistically insignificant). Altogether, the "None" option constitutes for 173, or 5.4%, of the total 3200 CBC responses which means that participants typically chose between the two profiles they were given.

6.3 Clarification Questions

The frequency of use results align with previous research in that the survey participants regularly use single-word search, sometimes use multi-word search and rarely use partial-word search. With the average search term being 1.49 terms [18], this implies users most often use single-word search. Moreover, these results tie in with the CBC results in that partial-word search is not used very often and it has the lowest relative importance.

Participants who reported using single-word search daily had higher relative importance for all of the search features (single-word search = 13.21 which is 12.6% higher, multi-word search = 14.59 which is 14.7% higher, partial-word search = 10.66

which is 8.8% higher). Participants who reported using multi-word search daily had an even higher relative importance for all of the search features (single-word search = 15.28 which is 24.4% higher, multi-word search = 15.24 which is 18.3% higher, partial-word search = 10.77 which is 9.7% higher). We see a dip in the relative importance of multi-word search for participants who reported using partial-word search daily but a higher relative importance for partial-word search (single-word search = 11.66 which is 0.9% higher, multi-word search = 8.28 which is 33.5% lower, partial-word search = 11.72 which is 20.6% higher). This reassures us that the participants were choosing profiles based on how they use their email.

Portability has a higher relative importance than the search features which make sense because users report using this feature more often. Participants who reported accessing messages on multiple devices daily had higher relative importance for the syncing old email to a new device feature (=14.67 which is 6.1% higher). Moreover, participants who said this feature is important also had higher relative importance for the portability feature (=14.42 which is 4.3% higher). Participants who mentioned that they would be unlikely to use an email service without message portability showed a slight increase in this feature's relative importance as well (=14.48 which is 4.7% higher).

The relative importance of storage aligns with the fact that 45% of the survey participants report a struggle with device storage. Intuitively, the local index email

approach uses the most device storage. These results suggest that users may shy away from local index email options to preserve space on their devices.

In line with the idea that while privacy is very important, users are not willing to give up valued features, are the results of the question asking if the participants would use an email service with maximum privacy but lacking various other features. The majority of participants would not use an email service that does not provide single-word search or is not portable. This implies that end-to-end encryption is not a viable secure email service solution for most users.

On the other hand, a small percentage of the participants would be willing to give up the usability features in order to obtain maximum privacy. For these participants, the relative importance of privacy was higher than the total relative importance of privacy. Table 25 displays these results.

To obtain maximum privacy…					
Feature	Total	No single-word search	No multi-word search	No partial-word search	Part w/ portability
Price	20.23	22.61	23.37	24.06	17.79
Single-word search	11.55	5.61	9.88	8.83	5.97
Multi-word search	12.45	15.44	10.91	11.84	12.92
Partial-word search	9.72	7.2	11.11	9.28	7.34
Privacy	18.7	24.28	21.02	22.93	27.47
Storage on your device	13.52	11.07	12.16	9.57	12.77
Syncing old email to a new device	13.83	13.8	11.56	13.5	15.73

Table 25: Comparison of participants who would accept maximum privacy over usability features (these values are not statistically significant)

An added benefit of asking clarification questions is that the participants are able to explain why they feel various features are important. Participants who reported single-word search as very important typically explained that with the number of emails in their accounts, using a search query is the only way to efficiently find a particular email. One participant said, “I recently bought a new computer but needed to find the email with my resume in it so I typed "Resume" and those emails were shown. If it wasn't for single-word search then I would have had to go through thousands of emails to find it.” Others admitted that they do not delete or organize their emails, so they need search capabilities to sort through the clutter.

Participants who ranked multi-word search as very important said that using a single-word search still returns too many irrelevant emails. Here a participant mentioned, “I don't want a vague search result. I have to include multiple words so that I can get the results that I want.”

Participants who reported partial-word search as very important explained that it takes less effort to develop a partial-word search query. Also, they noted that they cannot always remember a full keyword or correct spelling of a keyword, so partial-word search gives them a smaller subset of emails to sort through. Overall, participants noted that having search functionality saves them time and energy when trying to find an email.

Participants who see the importance of synching across multiple devices expressed that they are expected, both professionally and socially, to be responsive to their emails at all times, even when they cannot get to their desktop computer. A participant reported, "It would be a nightmare to keep up without everything syncing together."

6.4 Data Trends

An interesting trend, shown in Table 26, is between income and relative price importance. The relative importance of price tends to increase with income. This could indicate that users with higher income value money more and are less willing to pay for an email service.

Attribute	Total	Less than $25,000	$25,000 to $49,999	$50,000 to $74,999	$75,000 to $99,999	$100,000 to $149,999
Price	20.23	18	18.8	20.35	24.06	22.91
Single-word search	11.55	13.72	10.86	14.1	8.24	9.4
Multi-word search	12.45	11.35	12.91	10.79	14.63	13.55
Partial-word search	9.72	9.59	10.8	10.2	8.19	7.8
Privacy	18.7	19.04	20.05	17.61	17.87	18.73
Storage on your device	13.52	13.09	13.38	15.24	13.8	10.95
Syncing old email to a new device	13.83	15.21	13.2	11.7	13.21	16.66

Table 26: Trend between education and relative importance; statistically significant with $p < 0.01$ for price only

Table 27 shows the trend between having a score of 5.0 on the web skills scale and the relative importance of privacy. Those who reported having a full understanding of the various topics have a 15.1% higher relative importance for privacy than the average and a 20.4% higher relative importance for privacy than those who reported not having a full understanding of these topics.

Importance			
Attribute	Total	Full understanding of phishing & full understanding of spyware	Other
Price	20.23	17.22	21.72
Single-word search	11.55	12.67	11
Multi-word search	12.45	11.62	12.86
Partial-word search	9.72	8.99	10.08
Privacy	18.7	21.53	17.88
Storage on your device	13.52	12.86	13.84
Syncing old email to a new device	13.83	14.98	13.27

Table 27: Trend between having a full understanding of web skill security topics (phishing and spyware) and relative importance; statistically significant with $p < 0.01$

Tables 28 and 29 show the effect of price on privacy and email portability. These are the only statistically significant correlations observed in the data.

Table 28 presents that, other factors included, users selected a profile with no cost, but with standard privacy 9.9% more frequently than they selected the paid option with maximum privacy. This is much less than the other frequencies, which confirms that privacy is the most valuable feature.

Price x Privacy		
Price option	Privacy option	Count ratio
$0.00 (free)	Standard	0.514
$0.00 (free)	Extra	0.675
$0.00 (free)	Maximum	0.764
$1.99 (per month)	Standard	0.183
$1.99 (per month)	Extra	0.279
$1.99 (per month)	Maximum	0.415
Interaction Chi-Square		18.411

Table 28: Effect of price on the privacy feature

Table 29 presents that, other factors included, users selected a profile with no cost, but without portability 27.6% more frequently than they selected the paid option with single-word search.

Price x Syncing old email to a new device		
Price option	Sync option	Count ratio
$0.00 (free)	Yes	0.697
$0.00 (free)	No	0.603
$1.99 (per month)	Yes	0.354
$1.99 (per month)	No	0.233
Interaction Chi-Square		10.831

Table 29: Effect of price on the device syncing feature

Once again, these trends show that users are not willing to pay for their email service.

6.5 Future Work

This study provides a substantial amount of information about user requirements for email services. The logical next step would be to develop a secure email tool that meets the needs of the users. However, if we were to deploy the survey again, it would be interesting to see the weight of the features in the absence of price. It is difficult to disentangle the weight of price from the actual value of other features. While price is a feature that can alter a user's preference towards a product, this feature may need to be studied separately.

Another way to add value to this work would be to broaden the participant demographic. MTurk arguably provides the most reliable data for online surveys; however, it is likely that requiring participants to take the survey in a controlled environment would both cut back on inattentive responses and shed light on the relative importance of the features from a wider audience.

Lastly, with a bit of fine tuning this survey can be used by secure email service providers to suggest the option that would best fit each user. This would benefit the service providers because users would be more likely to use a service that meets their needs, rather than a service that is missing valued features.

Chapter 7: Conclusion

In this chapter, we revisit the research questions presented in the introduction.

- How can a better understanding of user preferences lead to usable email tools?

Knowing what the user needs or wants in an email service allows the designer to develop tools that meet these needs. If the user's needs are met and the service only adds benefit, users would have less reason not to adopt the tool. We cannot account for user's who wish not to migrate to a new or updated email tool for no reason other than that it is new.

- How can researchers determine the best design for secure email service platforms, while maximizing user privacy and considering user preferences?

Researchers could use a variation of this study to determine the needs of the client, individual or organization. The survey gives a clear view of the needs of the client and the HB analysis provides a concise output of the email service option best aligned with the client's preferences.

- What features should developers strongly consider when designing secure email services?

Developers have already made substantial progress in including the important expressiveness features in their tools. With the exception of classic E2EE, the secure email service options provide single-word search and multi-word search,

which are most important to users. Moreover, both local index options and cloud index options cover the great need for email portability. Based on the results, local index options could work for most users. Although device storage options are continually increasing, a large number of users still see device storage as a major concern. Until local index options consume lower local storage, searchable encryption may see broader adoption. For this reason, local index developers should work on reducing the size of the search index.

- Would users adapt to a secure email service missing features they consider very important?

Over 60% of users who reported having a full understanding of web security topics reported that they are unlikely to use an email service option that achieves maximum privacy, but lacks a feature of personal importance (71.2% for single word search, 68.8% for multi-word search, 68.2% for partial word search and 63.95% for portability averaged across the 2 web security questions). These results show how important it is to give users the features they desire in an email service option, as most users would not choose to adapt to a secure email service just for the sake of privacy.

References

[1] "Enrollment status of S&E graduate students, by field, citizenship, ethnicity, and race: 2014," The National Science Foundation, 2014. [Online]. Available: https://www.nsf.gov/statistics/2017/nsf17310/static/data/tab3-5.pdf.

[2] R. Rivest, A. Shamir and L. Adleman, "A Method for Obtaining Digital Signatures and Public-Key Cryptosystems," *Communications of the ACM*, pp. 120-126, 1978.

[3] M. D. Ryan, "Enhanced Certificate Transparency and End-to-end Encrypted Mail," in *NDSS Symposium 2014*, San Diego, 2014.

[4] S. Sheng, L. Broderick, C. A. Koranda and J. J. Hyland, "Why Johnny Still Can't Encrypt: Evaluating the Usability of Email Encryption Software," in *SOUPS 2006*, 2006.

[5] H. Orman, Encrypted Email: The History and Technology of Message Privacy, Woodland Hills, UT: Springer, 2015.

[6] S. Gaw, E. W. Felten and P. Fernandez-Kelly, "Secrecy, Flagging, and Paranoia: Adoption Criteria in Encrypted E-Mail," in *CHI 2006*, Montréal, 2006.

[7] A. De Luca, S. Das, M. Ortlieb, I. Ion and B. Laurie, "Expert and Non-Expert Attitudes towards (Secure) Instant Messaging," in *SOUPS 2016*, Santa Clara, 2016.

[8] N. Unger, S. Dechand, J. Bonneau, S. Fahl, H. Perl, I. Goldberg and M. Smith, "SoK: Secure Messaging," in *IEEE S&P 2015*, San Jose, California, 2015.

[9] D. X. Song, D. Wagner and A. Perrig, "Practical Techniques for Searches on Encrypted Data," in *IEEE S&P 2000*, Berkeley, California, 2000.

[10] M. S. Islam, M. Kuzu and M. Kantarcioglu, "Access Pattern Disclosure on Searchable Encryption: Ramification, Attack and Mitigation," in *NDSS Symposium 2012*, San Diego, California, 2012.

[11] M. Bellare, A. Boldyreva and A. O'Neill, "Deterministic and Efficiently Searchable Encryption," in *CRYPTO 2007*, Santa Barbara, California, 2007.

[12] E. Stefanov, C. Papamanthou and E. Shi, "Practical Dynamic Searchable Encryption with Small Leakage," in *NDSS Symposium 2014*, San Diego, California, 2014.

[13] B. R. Waters, D. Balfanz, G. Durfee and D. K. Smetters, "Building an Encrypted and Searchable Audit Log," in *NDSS 2004*, San Diego, California, 2004.

[14] B. Fuller, M. Varia, A. Yerukhimovich, E. Shen, A. Hamlin, V. Gadepally, R. Shay, J. D. Mitchell and R. K. Cunningham, "SoK: Cryptographically Protected Database Search," in *IEEE S&P 2017*, San Jose, CA, 2017.

[15] M. Harvey and D. Elsweiler, "Exploring Query Patterns in Email," in *ECIR 2012*, Barcelona, 2012.

[16] S. Whittaker, T. Matthews, J. A. Cerruti and J. C. Tang, "Am I wasting my time organizing email? A study of email refinding," in *CHI 2011*, Vancouver, BC, Canada, 2011.

[17] M. E. Cecchinato, A. Sellen, M. Shokouhi and G. Smyth, "Finding Email in a Multi-Account, Multi-Device World," in *CHI 2016*, San Jose, California, 2016.

[18] D. Carmel, L. Lewin-Eytan, A. Libov, Y. Maarek and A. Raviv, "The Demographics of Mail Search and their Application to Query Suggestion," in *WWW 2017*, Perth, Australia, 2017.

[19] Q. Ai, S. T. Dumais, N. Craswell and D. Liebling, "Characterizing Email Search using Large-scale Behavioral Logs and Surveys," in *WWW 2017*, Perth, Australia, 2017.

[20] J. Jordan, "53% of Emails Opened on Mobile; Outlook Opens Decrease 33%," Litmus Software, Inc., 2015.

[21] A. Oulasvirta and L. Sumari , "Mobile Kits and Laptop Trays: Managing Multiple Devices in Mobile Information Work," in *CHI 2007*, San Jose, California, 2007.

[22] P. E. Green and V. Srinivasan, "Conjoint Analysis in Marketing: New Developments with Implications for Research and Practice," *Journal of Marketing*, vol. 54, no. 4, pp. 3-19, 1990.

[23] H. Krasnova, T. Hildebrand and O. Guenther, "Investigating the Value of Privacy in Online Social Networks: Conjoint Analysis," in *ICIS 2009*, Phoenix, Arizona, 2009.

[24] Y. Pu and J. Grossklags, "Using Conjoint Analysis to Investigate the Value of Interdependent Privacy in Social App Adoption Scenarios," in *PETS 2016*, Darmstadt, Germany, 2016.

[25] D. Burda and F. Teuteberg, "Understanding the Benefit Structure of Cloud Storage as a Means of Personal Archiving - a Choice-Based Conjoint Analysis," in *ECIS 2014*, Tel Aviv, 2014.

[26] J. P. Turner, "University Preference: A Conjoint Analysis," Edith Cown University, Perth, Australia, 1999.

[27] L. Lockshin, J. Wade, F. d'Hauteville and J.-P. Perrouty, "Using simulations from discrete choice experiments to measure consumer sensitivity to brand, region, price, and awards in wine choice," *Food Quality and Preference*, vol. 17, no. 3-4, pp. 166-178, 2006.

[28] M. E. Kruk, M. Paczkowski, G. Mbaruku, H. d. Pinho and S. Galea, "Women's Preferences for Place of Delivery in Rural Tanzania: A Population-Based Discrete Choice Experiment," *Am J Public Health*, vol. 99, no. 9, pp. 1666-1672, 2009.

[29] W. Bai, C. Lynton, C. Papamanthou and M. L. Mazurek, "Understanding User Tradeoffs for Search in Encrypted Communication," in *EURO S&P 2018*, London, 2018.

[30] E. Stefanov, M. v. Dijk, E. Shi, T.-H. H. Chan, C. Fletcher, L. Ren, X. Yu and S. Devadas, "Path ORAM: an extremely simple oblivious RAM protocol," in *CCS 2013*, Berlin, Germany, 2013.

[31] M. S. Islam, M. Kuzu and M. Kantarcioglu, "Access Pattern disclosure on Searchable Encryption: Ramification, Attack and Mitigation," in *NDSS 2012*, San Diego, California, 2012.

[32] N. J. Taylor, A. R. Dennis and J. W. Cummings, "Situation normality and the shape of search: The effects of time delays and information presentation on search behavior," *Journal of the Association for Information Science and Technology*, vol. 64, no. 5, pp. 873-1089, 2013.

[33] R. Bost, "Σ*O*Φ*O*ς: Forward Secure Searchable Encryption," in *CCS 2016*, Vienna, Austria, 2016.

[34] M. Buhrmester, T. K. G. Kwang and S. D. Gosling, "Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data?," *Perspectives on Psychological Science*, vol. 6, no. 1, pp. 3-5, 2011.

[35] E. Peer, J. Vosgerau and A. Acquisti, "Reputation as a sufficient condition for data quality on Amazon Mechanical Turk," *Behavior Research Methods*, vol. 46, no. 4, pp. 1023-1031, 2014.

[36] B. K. Orme, "Assessing the Monetary Value of Attribute Levels with Conjoint Analysis: Warnings and Suggestions," Sawtooth Software, Inc, Sequim, WA , 2001.

[37] "The CBC System for Choice-Based Conjoint Analysis," Sawtooth Software, Inc., Orem, Utah, 2017.

[38] R. M. Johnson and B. K. Orne, "How Many Questions Should You Ask in Choice-Based Conjoint Studies?," Sawtooth Software, Inc., Sequim, Washington, 1996.

[39] M. Bech, T. Kjaer and J. Lauridsen, "Does the number of choice sets matter? Results from a web survey applying a discrete choice experiment," *Health Economics*, vol. 20, no. 3, 2011.

[40] A. Nuno and F. A. St. John, "How to ask sensitive questions in conservation: A review of specialized questioning techniques," *Biological Conservation*, vol. 189, pp. 5-15, 2015.

[41] D. J. Hauser and N. Schwarz, "Attentive Turkers: MTurk participants perform better on online attention checks than do subject pool participants," *Behavior Research Methods*, vol. 48, no. 1, pp. 400-407, 2016.

[42] E. Hargittai and Y. P. Hsieh, "Succinct Survey Measures of Web-Use Skills," *Social Science Computer Review*, vol. 30, no. 1, pp. 95-107, 2011.

[43] R. M. Johnson, "Understanding HB: An Intuitive Approach," Sawtooth Software, Inc., Sequim, Washington, 2000.

[44] Sawtooth Software, "CCEA Module," Sawtooth Software, [Online]. Available: https://www.sawtoothsoftware.com/products/advanced-analytical-tools/ccea.

[45] R. Kang, S. Brown, L. Dabbish and S. Kiesler, "Privacy Attitudes of Mechanical Turk Workers and the U.S. Public," in *SOUPS 2014*, Ottawa, Canada, 2014.

[46] R. Vetschera and G. Kainz, "Do Self-Reported Strategies Match Actual Behavior in a Social Preference Experiment?," *Group Decision and Negotiation,* vol. 22, no. 5, pp. 823-849, 2013.

[47] A. W. Meade and B. Craig , "Identifying Careless Responses in Survey Data," *Psychological Methods*, vol. 17, no. 13, pp. 437-455, 2012.

[48] R. Vetschera and G. Kainz, "Do Self-Reported Strategies Match Actual Behavior in a Social Preference Experiment?," *Group Decision and Negotiation,* vol. 22, no. 5, pp. 823-849, 2013.