Type, Talk, or Swype: Characterizing and comparing energy consumption of mobile input modalities

Author:
Jiang, F; Zarepour, E; Hassan, M; Seneviratne, A; Mohapatra, P

Publication details:
Pervasive and Mobile Computing
v. 26
pp. 57 - 70
1574-1192 (ISSN); 1873-1589 (ISSN)

Publication Date:
2015-01-01

Publisher DOI:
https://doi.org/10.1016/j.pmcj.2015.10.010

Type, Talk, or Swype: Characterizing and Comparing Energy Consumption of Mobile Input Modalities

Fangzhou Jiang[a,b], Eisa Zarepour[b], Mahbub Hassan[a,b], Aruna Seneviratne[a,b], Prasant Mohapatra[c]

[a] *National ICT Australia, Australia*
[b] *University of New South Wales, Australia*
[c] *University of California, Davis, United States*

Abstract

It is reported that mobile users spend most of their time on texting *SMS*, *Social Networking*, *Emailing*, or sending instant messaging (*IM*), all of which involve text input. There are three primary text input modalities, soft keyboard (*SK*), speech to text (*STT*) and *Swype*. Each one of them engages a different set of hardware and consequently consumes different amounts of battery energy. Using high-precision power measurement hardware and systematically taking into account the user context, we characterize and compare the energy consumption of these three input modalities. We find that the length of interaction, or the message length, determines the most energy efficient modality. For short interactions, less than 14-30 characters, *SK* is the most energy efficient. For longer interactions, however, *STT* significantly outperforms both *SK* and *Swype*. When message length distributions of popular text activities are considered, *STT* provides near optimal energy consumption without requiring the user to predict the message length and decide between *SK* and *STT*. In terms of battery life, the choice of input modality makes significant differences. If users always choose *SK* for all their text activities, they will consume nearly 50% of the phone battery each day. Choosing *STT* over *SK* can save 30% to 40% of the battery depending on the choice of *STT* software.

Keywords: mobile computing, text input modality, energy consumption efficiency, user context, recommendation

1. Introduction

Mobile device usage is becoming pervasive. A recent survey [1] has shown that the top activities of smart mobile device users are texting *SMS*, checking and sending Email, chatting and *Social Networking*. All these activities involve entering texts of different lengths. More detailed surveys show that a typical user sends about 110 text messages of an average size of 50 characters each [2, 3], 18 emails of 857 characters each on average [4, 5, 6, 7], 4 Tweets [8] and 42 instant messages [9]. Given these trends, it is clear that *text input is one of the major modes of interaction* for mobile device.

The users are also getting more concerned about the energy consumption of popular applications. A recent study of more than 9 million comments in Google play store has shown that more than 18% of all commented applications have negative comments with respect to their energy consumption [10]. Thus, with text input becoming the major mode of interaction, it pays to understand the energy consumption of this important mode of interaction.

Initially, text input was enabled on smart mobile devices via a soft keyboard (*SK*), e.g., typing on the touch screen keyboard. However with *SK*, the users in most cases, need to use both hands to type fast, which is difficult to do whilst "on the move". To address this, speech-to-text (*STT*) [11, 12] and *Swype*,

Email addresses: fangzhou.jiang@nicta.com.au (Fangzhou Jiang), ezarepour@cse.unsw.edu.au (Eisa Zarepour), mahbub@cse.unsw.edu.au (Mahbub Hassan), aruna.seneviratne@nicta.com.au (Aruna Seneviratne), prasant@cs.ucdavis.edu (Prasant Mohapatra)

which allows single-hand text input have appeared in the market. With *STT*, speech is captured on the mobile device and sent to a server for processing. With *Swype*, users simply swipe their finger from one letter toward the next of the intended word, and the mobile app attempts to predict the word. *Swype* is gaining popularity not only because it allows single-handed input, but also as it enables faster input compared to *SK* and is more discreet compared to *STT*.

Different users embrace different input modalities based on their habits, familiarity and convenience. The different text input modalities, however, use different hardware components and involve different amounts of processing on the mobile device. Consequently, they consume different amounts of energy for entering the same message. For example, *SK* predominantly uses the touch screen, while *STT* uses the microphone for recording and the communication interface for transmitting the speech sampled to a server and receiving the converted text from the server. Users are mostly unaware of the impact of these processing and communications requirements on the energy consumption of their text input. Given the large volume of text-based interactions, knowing the energy implications of the different input modalities will enable the users to make informed decision about which input modality to use to minimize their devices' energy consumption, especially when their device is low on power.

There has been significant work done in terms of optimizing the energy consumption of smart phones when it is being used for purposes such as video streaming [13], web-browsing [14], downloading content [15] and instant messaging [16]. It has also shown that using the least energy efficient application could potentially shorten the battery lifetime of a mobile device by a factor of 2.5 [16]. However, all these studies have focused on the use/running of the applications, and not on the user interactions. To the best of our knowledge, there has been no prior work in characterizing the energy consumption of user interaction with smart mobile devices. This paper addresses this through a comprehensive empirical study of energy consumption of the three widely used text input modalities, namely *SK*, *STT* and *Swype*.

Our contributions and findings can be summarized as follows:

- Using high-precision hardware, capable of capturing the true current drain from the device battery, we measure and characterize energy consumption of the three text input modalities.
- We show that for short interactions of less than 14-30 characters, *SK* is the most energy efficient, but *STT* outperforms *SK* for longer messages.
- The energy saving achieved with *STT* becomes more significant with increasing message lengths. If users always choose *SK* for all their text activities, they will consume nearly 50% of the phone battery each day. Choosing *STT* over *SK* can save 30% to 40% of the battery depending on the choice of *STT* software.
- We demonstrate that for *STT*, the application logic of whether to buffer speech samples before transmitting them to the remote server for analysis as opposed to streaming the speech has significant consequences for energy consumption.
- We further show that the "user style" and experience of using a given input modality has no tangible impact on the power consumption. In addition, the choice of input modality is independent of the device manufacturer and size of the device.

The rest of the paper is organized as follows. Section 2 summarizes related work. The detailed experimental setup and measurement methodologies are presented in Section 3. Results are presented and analyzed in Section 4, followed by some discussion in Section 5. Finally, Section 6 concludes the paper.

2. Related Work

Smart mobile device technology is improving rapidly with significant improvements in processing, storage and screen technology. These improvements are placing more and more demands on energy. However, battery technology is not keeping pace with these improvements and is unlikely to do so in the foreseeable future [17]. As a result, the research community has been investigating ways of minimizing the energy consumption of hardware exemplified by [18, 19, 20]. Similarly there has been considerable work done to make the applications more energy efficient, by reducing their interaction with hardware and communications. However, there has been limited work that have investigated the user interaction and the impact on energy.

Page [21] investigated the implications of typing using a soft keyboard, *ITU-T* numeric keypad, *Swype* and *Swiftkey* on six different smart phones. He concluded that, *Swype* and *Swiftkey* are the most effective and that they offer substantial benefits to users as typing speeds comparable to when using a common computer keyboard could be achieved. However, the study does not investigate the energy consumption of different input modalities. The focus of our work was to examine the energy consumption of the different text input modalities and provide a guide to the users as to which modality should be used to conserve energy.

Numerous groups have investigated energy consumption of mobile devices by examining the energy consumption of different hardware components of mobile devices and applications. Carroll and Heiser [22] presented the detailed breakdown of power consumption of mobile phone's main hardware components and developed a power model for smartphones. They investigated the energy usage and battery lifetime under the different usage patterns by analyzing the power consumption of the various components of a smartphone, and showed the most power hungry components. Yoon et al. [23] also used kernel activity monitoring as a way of determining the energy consumption of mobile applications and estimate the energy usage for online activities. These studies again focused on application behavior as opposed to user interaction.

Perrucci et al. [24] investigated the impact on energy consumption of a smart phone, when using different services such as data, cellular link services and mobile TV. They showed that for *SMS*, the energy consumption was dependent on the cellular network used. The overall finding was that *GSM* consumes less energy when compared to 3G (*UMTS*). While this finding influences our finding about the *STT* energy consumption, it does not directly address the impact of the input modalities on power consumption. Vergara et al. [16] studied the energy consumption of different instant messaging (*IM*) applications. They showed that short messages consume as much energy as longer messages and that it is possible to trade off latency for increased energy efficiency. Again Vergara et al. [25] showed that typing notifications result in almost a 100% increase in energy consumption. There are also a number of groups focusing on energy consumption of specific activities such as video streaming, web browsing and downloading. Trestian et al. analyzed the power consumption for video streaming using different wireless networks [13]. Their result showed the network load and signal quality together have a significant impact on energy consumption. Thiagarajan et al. [14] measured the energy consumption for web browsing using a similar measurement methodology to what is presented in this paper. They optimized the energy needed for web page downloading, rendering, and showed a modified Wikipedia mobile site which can reduce 30% of the energy cost. Energy consumed when downloading via different wireless networks (*WiFi*, 3G and Bluetooth) was also investigated by Kalic et al. [15]. They proposed an energy consumption model for each communication technology and showed that collaborative downloading could be used to lower the overall energy consumption. All these works, whilst relevant, do not address the impact of the input modalities on power consumption.

A preliminary version of this paper was presented at the conference [26], but this paper presents an extended analysis of the energy consumption. In particular, the conference paper only analyzed *SMS* texts and the mean message length of *SMS*. In this paper, we have analyzed three more popular text activities, *Email, Tweeter*, and *IM*. In addition, we have considered message length distributions (not just the mean length) for all four text activities. We have also considered a new input modality referred to as "*Oracle* modality" as a benchmark. The *Oracle* modality is capable of predicting the exact message length in the beginning and selects either *SK* or *STT* depending on the predicted message length. As a result of this extended analysis, it is possible to provide deeper insights to the energy consumption of different input modalities.

3. Measurement Methodology

Determining the impact of different input modalities on power consumption is difficult because of the large number of dependencies, especially the differences in user interaction styles and the context of use. To

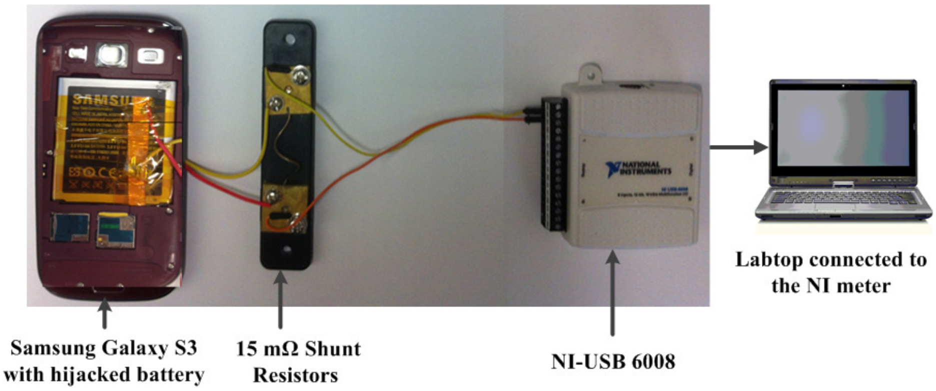


Figure 1: Power measurement setup

address differences in user interaction styles and the impact of context, two sets of experiments, referred to as primary and secondary experiments, were conducted. The primary experiments were aimed at identifying the key differences in power and energy consumption of the three input modalities, and the secondary experiments were aimed at identifying the dependency of the input modality power consumption on user contexts.

Although there exists several software power profilers for Android such as *BatteryManager* [27] and *CurrentWidget* [28] that could be used for power measurements, they only enable the measurement of power at fixed, system dependent intervals. For example, *Batterymanager* only gives the voltage readings whenever there is a percentage change in the battery level. Thus to measure power consumption at a finer granularity, for both sets of experiments, we used the set-up shown in Figure 1, which has also been used by others [14, 29, 30]. With this set-up, the smart mobile device battery is "hijacked" at one of its terminals, and connected in series with a $15\,m\Omega$ shunt resistor. Then a National Instrument (NI) NI-USB 6008 is used to sample the voltage drop, V, across the shunt resistor at 1 KHz and log the data on to a laptop computer. In addition, for each interaction, the start and end time was recorded and could be read directly from the voltage log file. Finally we used the standard equation of power, $P = V_b \times I_r$ to calculate the consumed power for each data point, where V_b is the battery voltage and I_r is the current through the shunt resistor. Then the average power consumption for a specific message input modality was computed as the mean value of all the calculated instantaneous power values during the interaction period, as determined by the logged start and stop times. The total energy consumed for a given input modality was calculated by multiplying average power by the interaction period.

3.1. Primary Experiments

These experiments were aimed at determining the power consumption of the input modality. Therefore, the experiments used a single fully charged ($\geq 95\%$) Samsung Galaxy S3 smart phone, connected to a 3G or WiFi network. Ten different users were asked to interact with the smart phone by entering the 7 text messages shown in Table 1, using each of the three input modalities. The lengths of the messages shown in Table 1 were chosen to be representative of the typical message lengths of text based interactions of smart mobile device users based on the distributions in [16].

To ensure that the power consumption was only due to the user inputs, all background processes on the smart phone were terminated via Android developer options, except for the application used for the experiment. The screen brightness is also set to a fixed level to eliminate any change during the experiment. Furthermore, after each interaction period, the battery level of the smart phone was checked and where necessary the smart phone was recharged to ensure that the battery level remained at or above 95%.

For *SK*, users entered messages using the Android's default Message App editor and default Samsung soft keyboard. Haptic feedback and auto-completion functions were disabled in the primary experiment to minimize the number of variables.

Table 1: Seven messages with different lengths

Message Length	Message Content
7	A phone
15	That was a test
27	These are few mobile phones
52	This is a test to investigate the energy consumption
79	This is a test to investigate the energy consumption of different mobile phones
102	This is a test to investigate the energy consumption of different mobile phones in different situation
202	This is a test to investigate the energy consumption of different mobile phones in different situation via variety of Android applications and games in various locations in university of New South Wales

For *STT*, the Galaxy S3's built-in *Google STT* application and the *STT* application available with the *Swype* application, namely *Dragon dictation* were used. They represented the only two *STT* applications available [1]. With Google's *STT* application, each phrase of speech is recorded and then streamed to a Google sever for conversion from speech to text. Once the converted text from the server is received, it is displayed on the screen. *Dragon dictation* operates in a similar manner, except that it records the speech for given period of time and sends it to a cloud based server for the speech to text conversion. The difference is that *Dragon dictation* application always tries to record for as long as possible, up to 100 seconds of recording, before sending the whole recorded segment to the server. Because the streaming nature of *Google's STT* application, the communication modules on the mobile device is kept in the active state [31], it consumes more power than *Dragon Dictation*, but has less latency which is reported to result in better user *QoE* as the users can actually see what is being typed in real time. In addition, we also find that, *Google STT* utilizes location service by default, in contrast, *Dragon dictation* does not. This also leads to extra power consumption. As it is the default configuration, we decided to keep the location service on for the primary experiments.

Swype keyboard application was used for the *Swype* experiments. This involved the users simply tracing the characters of a word with their fingers, and the software predicting the word and displaying it on the screen. All volunteers were allowed time to become familiar, if they had not used *Swype* before to minimize the user biases.

In the primary experiments, users are asked not to correct any mistakes they make, and record the mistakes once they have finished. The details of error characteristics and error corrected energy consumption is shown in the secondary experiments.

3.2. Secondary Experiments

The objective of these experiments were to investigate the impact of user context on the power consumption of the three input modalities. Thus the experiments involved a single user interacting with three devices, two smart phones (Samsung Galaxy S3, S4) and a tablet (Google Nexus 7) using the same messages used in the primary experiments.

The contexts of different devices operating at two different battery charge levels of 95%, 30% and connecting to two different networks (3G/WiFi) in the case of a smart phone were evaluated. Additionally, the overhead of utilizing *GPS* was investigated. During these experiments, to minimize the network connectivity variations, the experiments were repeated in three different locations, namely inside a research lab in the city center, inside a residential apartment in a suburb, and inside a student laboratory.

In order to investigate the impact of "user typing style" on *SK* power consumption, we developed an Android application that logged the touch down/up time, holding time, pressure and the size of the touch [2].

[1] all other applications use the *Google STT* engine.

[2] pressure could not be recorded for Samsung S3 because it uses a capacitive screen.

Table 2: Comparing power consumption of text related activities with other non-text online/offline activities via S3

Online Activities	Power consumption	Offline Activities	Power consumption
†Email via browser	2.38W	**Taking photos**	2.8W
Video Streaming	2.15W	**Recording video**	2.57W
Online Games	2.15W	**Playing games**	2.01W
News/Weather	1.59W	**Call facilities**	1.47W
†Web Surfing	1.57W	**†Texting/IM**	0.98W
†Email via App	1.48W	**Listening to Music**	0.86W
	Normal Usage	0.68W	
	Idle	0.48W	

† Activities involve text input.

The impact of haptic feedback function was also investigated by comparing the energy consumption of completing the same message whilst the function is on and off. All experiments are repeated three times and the average was used for analysis.

4. Results

In this section, we report the outcome from both the primary and the secondary experiments. Results from the primary experiments are used to identify the key differences in power and energy consumption of the three input modalities. The secondary experimental data are used to verify whether these key differences are dependent on user contexts.

Before we present these results, we compare the power consumption of some major online and offline activities presented in Table 2. The baseline power consumption, e.g., normal usage and idle are also shown. In the idle state, the phone is in airplane mode with all communication modules disabled. The screen is turned on and no background applications are running. Normal usage represents the state when all communication modules are enabled without any background applications. The phone consumes 0.68 W on average in normal usage and 0.48 W when idle.

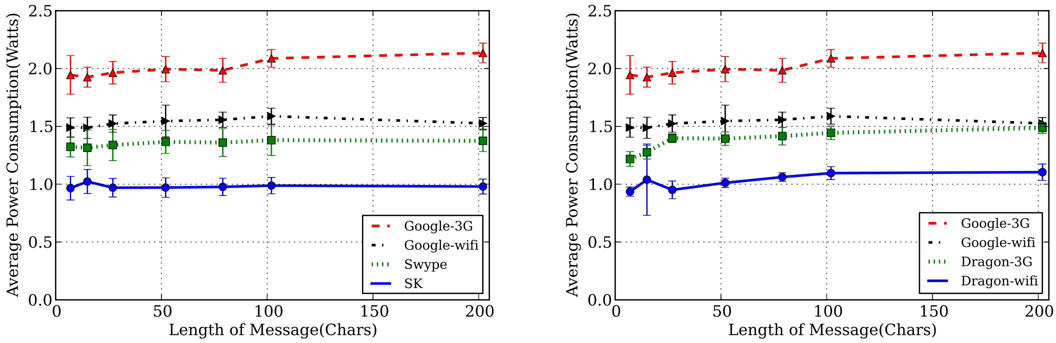
As can be seen, online activities consume as much power as offline activities. For example, playing offline games consume 2.01 W, which consumes more power than web browsing (1.57 W). In particular, checking/sending emails via browser could consume as much as 2.38 W, while the same activity via an *App* only consumes 1.48 W. Note that both of the two activities measures the power consumption of sending/receiving emails instead of inputting text. In addition, text related activities consume as much power as non-text activities, which motivates the need to better understand the energy consumption of different text input modalities and identify opportunities to reduce it.

4.1. Primary Experiment Results

In this section, we first analyze *power consumption* of different input modalities. Energy consumption of a message depends not only on the power consumption, but also on the time taken to enter the message. We therefore also study the input *speed* of different modalities, e.g., how fast users can enter text using different modalities, which then allows us to study energy consumption of a given modality. Finally, we seek to identify the *optimal modality* that would *minimize energy consumption* of text input by analyzing message patterns of four popular text activities, Short Message Service (*SMS*), Tweets, Email, and Instant Messaging (*IM*). To minimize the number of dependencies, error correction is not performed in these experiments, however, is recorded and further analyzed in the secondary experiments.

4.1.1. Power Consumption

Figure 2 plots the average power consumption (averaged over 10 users) of different input modalities as a function of message length. First, we find that power consumption is independent of message length for all three input modalities. Second, different modalities consume different amounts of power. Table 3 lists power



(a) *Google STT*, *Swype* and *SK* power consumption

(b) *Dragon dictation* vs *Google STT*

Figure 2: Power consumption comparison of input modalities

Table 3: Power consumption of different input modalities

Input Modality	Power Consumption
SK	0.98 W
STT-D WiFi	1.03 W
Swype	1.35 W
STT-D 3G	1.38 W
STT-G WiFi	1.55 W
STT-G 3G	2.01 W

Table 4: Input speeds for different modalities

Modality	Completion Time Slope(m)	Input Speed(char/sec)
SK	0.457	2.19
Swype	0.422	2.37
STT	0.074	13.51

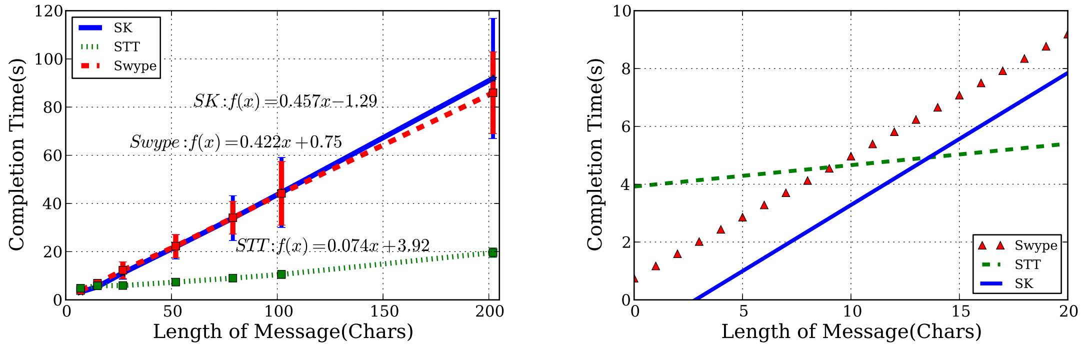
consumption of different modalities in ascending order. We see that SK consumes the least energy ($\sim 1W$). *Swype* consumes 30% more power than *SK*. *STT* also consumes more power than SK with majority of the *STT* versions consuming the most energy. For example, *Google STT* with 3G (*STT-G* 3G) consumes 2W, which is twice as much consumed by SK. The reason that *Google STT* consumes more power than *Dragon STT* is not only because of the difference in speech sampling window as discussed, but also because *Google STT* utilizes *GPS*. This will be further discussed in the secondary experiments. Similarly, *STT* consumes more when used with *3G* compared to when used with *WiFi* because *3G* interface is known to consume more power than *WiFi* [32].

At this point, it is important to remember that power consumption reflects only part of the overall energy consumption. Speed of message input is also an important factor. For example, if two input modalities consume the same amount of power, but one allows the user to complete the message quicker, then it will consume less energy than the other. Therefore, it is necessary to take into account the *speed of completion of the different input modalities* before we can compare their energy consumption.

4.1.2. Speed of Input Modalities

For the three input modalities, *SK*, *Swype*, and *STT*, Figure 3a shows message completion times as a function of message length. All versions of *STT* had the same completion time curve, hence we show a single curve for *STT*. From these results, it is possible to make several observations. First, for all three modalities, completion time increases *linearly* as a function of message length. The speed of input therefore can be directly derived as $\frac{1}{m}$, where m is the slope of the line.

Second, *SK* and *Swype* have similar speeds (around 2 char/sec), but *STT* is 6 times faster (see Table



(a) Completion time (b) Zoom in

Figure 3: Message completion time comparison of three input modalities

4 for exact speeds). This means that *Swype* cannot outperform *SK* because it consumes 30% more power as discovered earlier. *STT*, however, can potentially outperform *SK* due to its significantly faster speed. Note that the speed of 2.19 char/sec that we derived from our SK measurements (averaged over 10 users) is within the theoretical typing speed upper bound of 2.51, which is based on Fitt's law [33].

The third observation relates to the 'setup' time. Because *STT* needs to transmit the message to a remote server before the text can be detected and entered, there is a higher setup delay compared to *SK*, which processes the input locally. Therefore, for short messages, *SK* enjoys lower completion times. This initial advantage for *SK*, however, starts to erode with increasing message size. As can be seen in Figure 3b, although *STT* starts with higher completion times at the beginning, it outperforms SK after about 13 characters. *STT* therefore has the potential to consume less energy than SK despite consuming significantly more power. In the following section, we will compare these two input modalities with regards to their energy consumption.

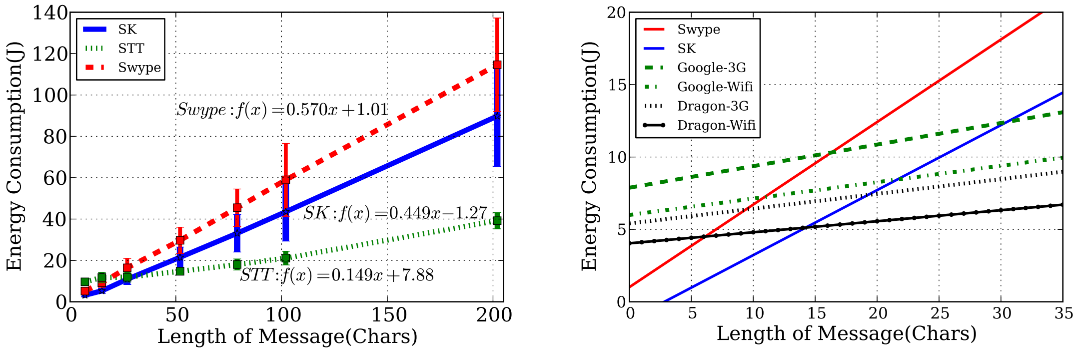
4.1.3. Energy Consumption

For the three input modalities, Figure 4a shows their energy consumption derived as completion time multiplied by power consumption. It shows that STT is more energy efficient than *SK* for longer messages despite its higher power consumption, as the completion time decreases with longer messages. *Google STT* with 3G becomes more energy efficient than *SK* for messages longer than 30 characters, but *Dragon STT* with WiFi outperforms SK only after 14 chars due its low power consumption. This finding demonstrates that relative superiority of one input modality over another with respect to energy consumption is heavily dependent on the length of message. In the following section, we provide a deeper analysis of energy consumption based on empirical distributions of message lengths for popular text activities.

4.1.4. Optimal Modality Selection

In this section, we seek to identify the input modality or the combination of modalities which would minimize daily battery energy consumption for a typical user. Because *Swype* was not found to be as energy efficient as *SK* or *STT*, we do not consider *Swype* as a candidate modality to minimize energy consumption.

In the previous section, we have found that for short messages, e.g., shorter than 14 (*STT-D*) or 30 (*STT-G*) characters (lower and upper bound of STT), *STT* consumes more energy than *SK*, but the situation is reversed for longer messages. It is therefore difficult to answer the question which choice of input modality would minimize energy consumption. For example, if the messages are predominantly smaller than 14, then SK is the optimal choice. On the other hand, if messages are generally very large, then *STT* is a clear winner. The answer therefore lies with the *distribution of message lengths*. In this section, we consider empirical distributions of different text activities to find the optimal input modality that minimizes energy consumption.



(a) Overall (b) Zoom in

Figure 4: Energy consumption comparison of three input modalities

Table 5: Message length distribution of different text activities

Text Activity	Distribution and Parameters	Maximum Length
SMS	Normal ($\mu = 50.9, \sigma = 46.2$) [3]	160 characters
Tweet	Log-Normal ($\mu = 0.55, \sigma = 4.3$) [35]	140 characters
Email	Normal ($\mu = 857, \sigma = 2346$) [6]	NA
IM	Normal ($\mu = 26, \sigma = 46.2$) [16]	NA

Before we compare *SK* against *STT*, we would like to consider a third option that we will call *ORACLE*. In *ORACLE*, the user can predict the exact message length before he starts entering the text. For *ORACLE-D*, he chooses *SK* if the message is predicted to be less than 14 characters, and *STT-D* otherwise. Similarly, for *ORACLE-G*, he chooses *SK* if the message is predicted to be less than 30 characters, and *STT-G* otherwise. Clearly, *ORACLE* is not achievable in practice, but it serves as the theoretical lower bound for energy consumption.

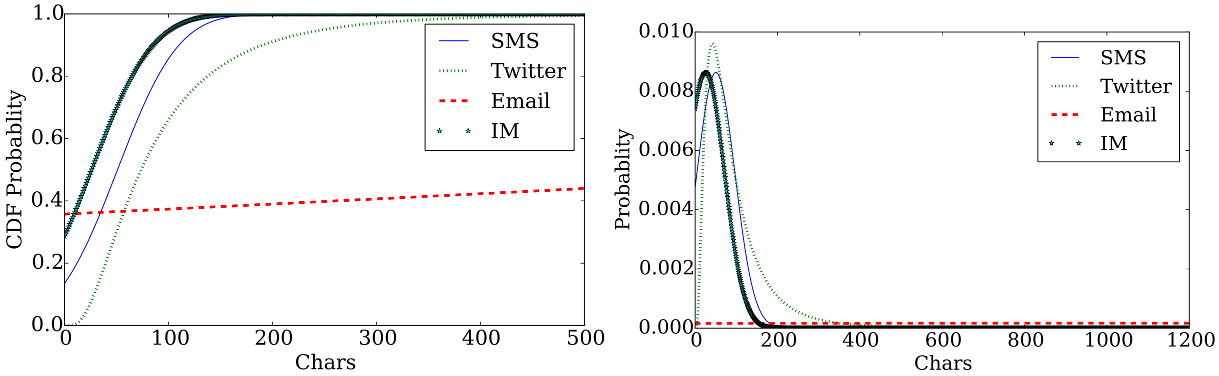
As mentioned earlier, we will consider distribution of message lengths of *popular* text activities to study the energy consumption of different input modalities. One way to measure popularity of mobile phone applications is by measuring the time spent with the application. An application is more popular if the users spend more time with it, and vice versa. A recent Nielsen report [34] shows that *SMS*, *Social Networking*, such as *Twitter*, *Email*, and *IM* are among the top 6 activities that users spend most of their times with. The message length distributions of these four text activities along with the references are shown in Table 5.

Using the values of Table 5, we plot the PDFs and CDFs of these four text activities in Figure 5. It could be seen that *Email* inputs are significantly larger than the other ones, which is expected due to the nature of the application. An interesting observation is that although Twitter messages are physically limited to 140 characters, the CDF does not cover 100% at 140. This is because users sometimes send long messages using back-to-back 'tweets', which are counted as a single Twitter message [3].

Next, the *expected energy consumption* of each message is derived using the empirical message length distributions of Table 5. For *SK* and *STT*, we obtain:

$$e^{modality}_{activity} = \sum_{n=1}^{N} \{CDF_{activity}(n+1) - CDF_{activity}(n)\} \times y_{modality}(n) \tag{1}$$

where $modality \in \{SK, STT\text{-}D, STT\text{-}G\}$ and $activity \in \{SMS, Twitter, Email, IM\}$. $y_{modality}(n)$ is the linear energy consumption functions for specific modalities as shown in Figure 4b. A large value for N ($N = 20000$) is chosen to cover large message lengths possible in Email.



(a) CDF

(b) PDF

Figure 5: Message length distributions of different text activities.

For ORACLE-D, *expected energy consumption* is computed as:

$$\begin{aligned} e^{\text{ORACLE-D}}_{activity} &= \sum_{n=1}^{14} \{CDF_{activity}(n+1) - CDF_{activity}(n)\} \times y_{\text{SK}}(n) + \\ & \sum_{n=15}^{N} \{CDF_{activity}(n+1) - CDF_{activity}(n)\} \times y_{\text{STT-D}}(n) \end{aligned} \tag{2}$$

Similarly, *expected energy consumption* for ORACLE-G is computed as:

$$\begin{aligned} e^{\text{ORACLE-G}}_{activity} &= \sum_{n=1}^{30} \{CDF_{activity}(n+1) - CDF_{activity}(n)\} \times y_{\text{SK}}(n) + \\ & \sum_{n=31}^{N} \{CDF_{activity}(n+1) - CDF_{activity}(n)\} \times y_{\text{STT-G}}(n) \end{aligned} \tag{3}$$

From *expected energy consumption* we compute *daily energy consumption* as:

$$E^{modality}_{activity} = \alpha \times e^{modality}_{activity} \tag{4}$$

where α is the average number of messages entered daily for a given activity.

Table 6 shows the daily energy consumed ($E^{modality}_{activity}$) by different text activities and input modalities. The numbers in the parenthesis next to the activity names represent the average number of messages entered for each activity per day. For example, 110 *SMS*, 4 *Tweets*, 18 *Emails*, and 42 *IMs* are sent each day on average [2, 4, 5, 8, 9]. The last column shows the total energy consumption per day from all text activities. We make the following observations:

- If users always choose SK for all their text activities, they will consume a total of 14,837 J each day, which is nearly 50% of the phone battery capacity (Samsung S3 battery has a capacity of 28,728 J).
- Choosing *STT* over *SK* will drastically reduce daily energy consumption. *STT-G* would save 8,738 J or 30% of the battery and *STT-D* would save 11,715 J or 40% of the battery.
- Surprisingly, *ORACLE* cannot save much beyond *STT*, i.e. *STT* performs close to the theoretical optimal. There are two factors that contribute to this outcome. First, as we can see in the CDF graphs (Figure 5a), only a small fraction of the messages are shorter than 14 characters. Second, there are many long emails, which consume significant energy when *SK* is used, but it is precisely the large length that helps *STT* to save energy. This is an encouraging finding, because it means that one can achieve near optimal energy consumption by simply using *STT* for all activities.

Table 6: Daily energy consumption of different activities using different input modalities

Input Modality	SMS (110)	Twitter (4)	Email (18)	IM (42)	Combined
SK	2,528J	186J	11,509J	614J	14,837J
STT-G	1,628J	98J	3,915J	458J	6,099J
ORACLE-G	1,536J	97J	3,915J	412J	5,960J
STT-D	834J	50J	2,003J	235J	3,122J
ORACLE-D	810J	50J	2,003J	221J	3,084J

Table 7: Energy Consumption (Jouls/char) of the Smartphone and Tablet

Input Mode	Samsung S3	Samsung S4	Tablet Nexus7
STT-G-WiFi	0.11	0.12	0.27
STT-D-WiFi	0.08	0.10	0.19
SK	0.45	0.59	0.87
Swype	0.57	0.61	1.02

Table 8: Power consumption of location service

Input Mode	S3	S4	GPS consumption	
			S3	S4
Google-Wifi-GPS-on	$1.53W$	$1.65W$		
Google-Wifi-GPS-off	$1.29W$	$1.47W$	$0.24W$	$0.18W$
Google-3G-GPS-on	$2.03W$	$2.10W$		
Google-3G-GPS-off	$1.73W$	$1.95W$	$0.30W$	$0.15W$

4.2. Secondary Experiment Results

4.2.1. Devices, Networks and Battery Charge Level

The secondary experiments used the Samsung Galaxy S3 used in the primary experiments, a Samsung Galaxy S4 and a Google Nexus 7 tablet. They are connected to *WiFi* and/or 3G cellular networks. The Samsung Galaxy S4 represents devices with more powerful hardware and Nexus 7 tablet represents devices with bigger screens and larger batteries. Table 7 presents the energy consumption rates of the three devices for the all input modalities. It shows that the tablet consumes approximately twice as much energy as the S3 per character for all text input modalities. In addition, S4 consumes slightly more energy than S3 mainly due to a higher resolution screen and a faster processor.

Figure 6 shows the energy consumed by the three devices as a percentage of the device battery, when used for texting by the 13-17 age group. The results show that all devices display similar characteristics despite the tablet consuming twice as much energy per character and the tablet battery having double the capacity (16Wh/57600J) of the S3 battery (7.98Wh/28728J). The S4 battery (9.88Wh/35568J) is around 25% larger than S3's, which is however not always enough to offset the extra power consumption of hardware.

In the case of STT, both smart phones display more than double the energy efficiency of the tablet and in the case of *SK*, the tablet performs a bit better battery percentage wise. We speculate that this due the smart phones having more energy efficient communications hardware than tablet. Since the percentage increase in energy consumption is similar for all the three input modalities across the three devices, the main observation of previous subsection, namely the choice of different input modalities is device independent. In other word, the battery life of different devices are similarly affected by choice of input modalities.

As STT is dependent on sending data to server for analysis, its energy consumption will be influenced by the network connectivity. To assess the impact, we measured the power consumption of STT when the smart phone was connected to a *3G* cellular network, and a *WiFi* network, in three different locations, namely inside a research laboratory, a residential apartment, and inside a student laboratory at a University. The mean power consumption of STT-G in the three locations were all approximately 2 W with a standard deviation under 0.1 W. Hence the results of previous subsection is also location independent.

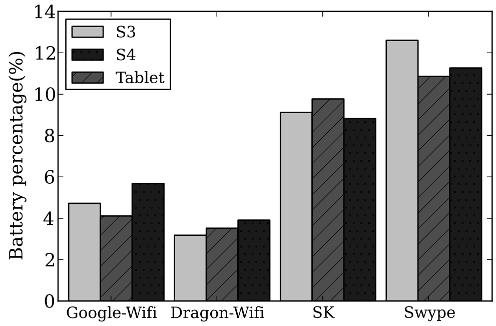


Figure 6: Battery level consumption for phone and tablet

When the experiment was repeated at a battery charge level of 30%, the current drain is increased by 5%-10% comparing to when the smart phone was fully charged ($\geq$ 95%). This we believe is due to the discharge behavior of chemical batteries as explained in [36]. The regulator used for power management increases the current drain to compensate for battery voltage drop. However, this again does not affect the main observation as the power consumption is not dependent on the battery charge level.

Thus the primary results (choice of input modalities) will remain the same, regardless of the screen size, processing power and battery charge level and will only be influenced by the network type, the application and usage.

4.2.2. Location service

Google utilizes location service for *STT* conversion, it is also interesting to know the portion of extra power consumption due to location service and speech streaming respectively. We found almost no difference in terms of power consumption for *STT-D* while keeping the *GPS* module on and off. The *GPS* symbol on the phone also indicates *STT-G* uses *GPS* as oppose to *STT-D*. Table 8 shows the power consumption of *STT-G* for the two difference smart phone under different network type and *GPS* status. The net power consumption used on location service is deduced, which shows an $0.2W$ is drawn from the *GPS* module on average.

However, the status of location service again would not affect observations in primary experiment. Both the results of *STT-G* and *STT-D* are presented, representing upper and lower boundary of *STT* energy consumption. Only slight change in the intersection points is observed, the main trend remains the same.

4.2.3. "User Typing Style" Analysis

For *SK*, the energy consumption could be influenced by the user "typing style". To investigate this we analyzed effect of touch size and touch duration each key press by developing a simple application which measured the touch time and duration.

Figure 7 shows the touch time of users during the experiments, which is done by using Gaussian kernel density estimation. It shows that the difference between the mean touch times of users is approximately $55\,ms$, and 80% of all touches has a touch duration under $90\,ms$. It is clear that the touch time varies from user to user and shows distinct touch styles each user has. Users type faster/touch lightly will have the touch distribution is pushed to the left hand-side, while users having a more preferable touch time will have the distribution more concentrated as shown in the graph. Touch sizes were found to be similar among all the users, with a mean touch size to be $0.04\,cm^2$ and standard deviation of approximately $0.01\,cm^2$. When individual touch is analyzed separately, the power consumption of each touch is found to be independent on both touch size/pressure [3]. However, It is also observed that energy consumption of each touch is positively correlated with the touch size/pressure. This is mainly due to the fact that touch size/pressure is correlated with touch time, where a larger touch size or stronger pressure is often associated with longer touch time.

[3]The pressure results are taken from Nexus 7 tablet, which has a resistive screen

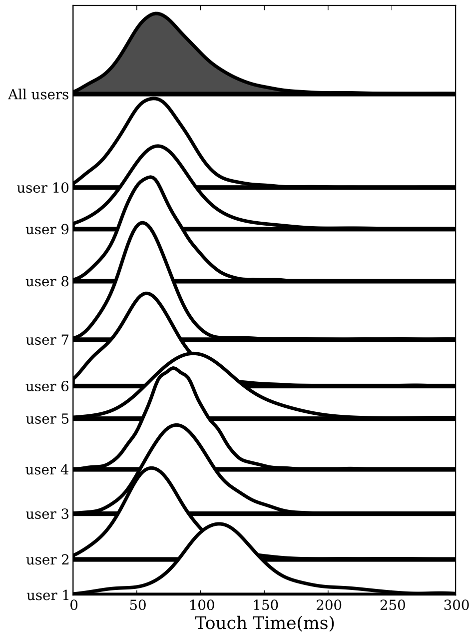


Figure 7: Probability density function of the touch time for users

This analysis shows that typing style, namely the speed and the weight of touch, result in variation of power consumption between $\pm 6\%$ of 1 W. Thus, the different "typing styles" of users do not have tangible impact on the power consumption of *SK*.

4.2.4. Impact of Haptic feedback

The same message was input three times under two scenarios: haptic feedback on and off. The average power consumption for turning haptic feedback on was 0.973 W, while keeping haptic feedback off consumes an average 0.966 W. As a result, hapatic feedback has no significant impact on the power consumption of SK as well.

4.2.5. Error Characteristics

SK, *Swype* and *STT* display different error characteristics. The *SK* errors are random and evenly distributed (trickle errors), where as the *Swype* and *STT* be word specific and thus tend to occur in groups (burst errors). The two *STT* engines displayed different error behavior. *STT-G* showed higher accuracy (91.69%) when compared to *STT-D* (77.11%) for the set of experiments which consisted of the 7 inputs as shown in Figure 8. The accuracy was shown to be lower than when using *SK* (96.47%) or *Swype* (92.38%). Also, *Swype* error rate increased with the length of interaction.

There have been significant work done in terms of text entry error and correction models [37, 38, 39]. Users correct errors differently depending on whether they are using *STT/Swype* or *SK*, thus we use different error correction model for different input modalities. Whenever a user makes a mistake when using *SK*, it is generally a single character. Thus the error is corrected by deleting the erroneous character and enter the correct character. Therefore with *SK*, each error results in 2 key strokes. As a result, for a message of length L chars, if one assumes an error rate of $e\%$ and that the user will only need to delete the wrong character once, the final length of character input interaction would be $L \times (1 + 2e)\%$. Because overall user's input speed when using *SK* is constant, i.e. that the completion time is linear, error corrected completion time can be derived, the energy consumption calculated.

On the other hand, for *Swype* and *STT*, the prediction engine will underline the words that could be in error. Assuming users take t seconds to press the word and choose the right word in a list of words with

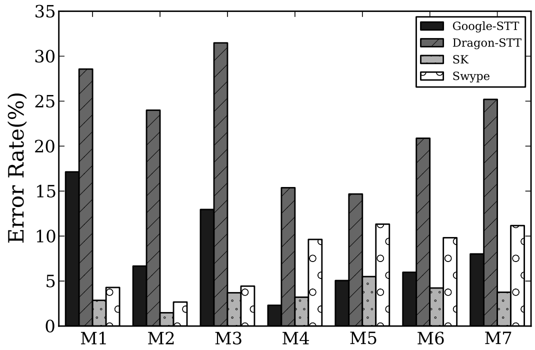
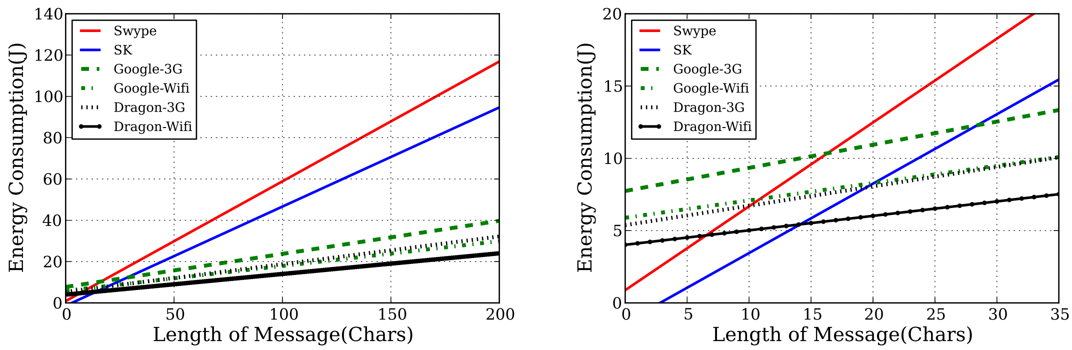


Figure 8: Error rate for all input modalities



(a) Overall

(b) Zoom in

Figure 9: Error Corrected Energy consumption comparison of three input modalities

similar pronunciation, a message length L characters and an error rate $e\%$, the average number of words that need to be corrected can be determined. In turn the extra time, and hence the energy taken to correct all the errors could be determined. Finally, t can be estimated by adding mean press duration from Figure 7 and thinking/reaction time, where a total number of $500ms$ is used in the calculation. The results obtained using the above two methods is shown in Figure 9a. Comparing Figure 9a with Figure 4a, the error rates of the different text input modalities do to have a significant impact, and therefore the findings of the primary experiments still hold.

4.2.6. Optimal Modality Selection with Error Correction

It can be seen in the error corrected energy consumption Figure 9a, the intersections where *STT-D* and *STT-G* consumes more energy than *SK* are 14 chars and 28 chars. The energy consumption fitting functions change very slightly, therefore we expect no tangible changes on the conclusions we draw in the primary experiments.

5. Discussion

5.1. Observations

This study has shown that overall, of the three text input modalities that are commonly used, the *SK* has the lowest energy consumption for short interactions. For longer interactions, the *STT* has the lowest energy consumption. *Swype* on average is the least energy efficient. The results also show that, these findings hold true regardless of the variations in user usage and speaking styles, the type of access network being used, and

Table 9: Characteristics of Different Input Modalities

Input Modes	**Accuracy**	**Convenience**	**Privacy**	**Speed**	**Energy Consumption**	
					Short	*long*
SK	Highest	Low	High	Fast	Lowest	Low
STT	Lowest	High	Low	Fastest	High	Lowest
Swype	Medium	High	High	Slower	Low	High

the type of device that is being used. There is also higher potential for STT to have better efficiency gains than SK as the technology improves this will make STT the most efficient form of interaction, except for very short interactions (less than 5 characters). Also, it is clear from the finding that, the streaming *Google* STT as opposed to "batch" processing *Dragon* STT models have significant energy implications because of the energy overheads of keeping the communication hardware of the mobile device in an active state. The obvious solution is to consider a hybrid approach where "adaptive bundling" is used. This warrants further investigation. Furthermore, *Google* STT utilizes GPS module resulting in a 10-20% increased power usage on average, which could be further optimized.

When the length distributions of different input categories are taken into consideration, overall the expected energy consumption of STT is always lower than *SK*. For inputting emails in mobile phone, using STT could at least reduce the energy consumption by 60%. However, there is little gain for always switching to the optimal modality. On the other hand, it would be worthwhile for users to switch between *SK* and STT based on the input length of *Instant Messaging* for an extra 7.5% reduction in energy consumption.

With the current availability of STT and *Swype* applications and the trends in text based interactions of smart mobile device users, such as messaging and social media interactions, there will be clear choice for users from solely an energy point of view, as most of these interaction will involve pressing a button, or swiping the screen. However, there are many other factors that will influence users choice of the input modality. Table 9 provides a comparison of what we believe will be the most important of these factors, and provides a subjective assessment of the benefits of *SK*, STT, and *Swype* with respect to these factors. When all the factors are taken into account, none of the three input modalities standout as the obvious choice. Therefore it is necessary to develop a recommendation system that takes into account these factors and acts as guide for the users as currently most users are unaware of the implications specially with respect to their battery usage. As a simple change, we propose to set the default input modality for long input task, i.e. Email, as STT.

5.2. Limitations

There are potentially a number of limitations of the experiment that were carried out. Firstly, we only considered the input modalities with respect to English. This presents the best case scenario, as STT and *Swype* engines are optimized for English. Although it is possible that other languages provide different results, we do not expect major impact on our results by considering English only as the language will equally affect all input modalities. Second, the sample sizes and the user population that were used was small. This was necessary because we needed to use the experimental setup discussed in Section 3 which required the device battery to be "hijacked" to get fine grained energy measurements. We attempted to mitigate this by having users of different nationalities (4 nationalities) and range of age groups (20-50 years) who were regular smart mobile device users. Further we used inputs that are representative of the type of interactions these users would have, from published data. Therefore, despite the sample being small, we do believe that our results are representative. Use of a higher number of users we believe would not lead to significantly different results. Third we only used three devices with a single operating system, *Android*. Despite the differences in the three devices, we could not see any indication that our results were device dependent. As for the operating system, it was not possible to carry out the same set of experiments on *iOS*. We believe this is not a real limitation as the overall findings will be applicable across platforms as the fundamental reasons for the differences stamp from the users, applications and the use of the different hardware components of the smart mobile device. Finally, although the error rates when using STT tend to be higher than when using *SK*, the methodology used provides a fair comparison for two primary reasons. (a) The accuracy of STT is improving and (b) the power consumption of STT at the longer lengths is significantly lower than that of *SK*. Therefore, overall STT will be the most energy efficient at longer lengths (greater than 30 characters).

6. Conclusion

For mobile devices, text input has become one of the major modes of interaction. Consequently, there is urgency to better understand the energy consumption dynamics of this important interaction. Since text can be entered using different input modalities, any energy consumption study for text interaction must be done with regards to these modalities. In this paper, we have studied energy consumption dynamics of three primary input modalities, *SK*, *Swype*, and STT. Our findings suggest that choice of input modality has a major influence on energy consumption of text input. By choosing STT as the universal modality for all types of text inputs, a typical user can save 30-40% of the battery depending on the choice of STT software. STT software that buffers speech samples at the device and sends them in bulk to a remote server for conversion to text can save significant energy compared to those that stream the speech to the server.

7. References

[1] O. P. Assciations, A portrait of today's smartphone user.

[2] N. Company, U.S. Teen mobile report calling yesterday, texting today, using apps tomorrow.

[3] A. Battestini, Vidya Setlur, T. Sohn, A Large Scale Study of Text Messaging Use, in: Proceedings of ACM MobileHCI'10, Lisbon, Portugal, 2010.

[4] S. Radicati, J. Levenstein, Email Statistics Report, 2014-2018 44 (0) (2014) 2014–2018.

[5] J. JORDAN, 53% of Emails Opened on Mobile; Outlook Opens Decrease 33%, Tech. rep. (2015). URL `https://litmus.com/blog/53-of-emails-opened-on-mobile-outlook-opens-decrease-33`

[6] F. Kooti, L. M. Aiello, M. Grbovic, K. Lerman, A. Mantrach, Evolution of conversations in the age of email overload, in: Proceedings of the 24th International Conference on World Wide Web, International World Wide Web Conferences Steering Committee, 2015, pp. 603–613.

[7] N. Arab-Moghaddam, M. Senechal, Orthographic and phonological processing skills in reading and spelling in Persian/English bilinguals, International Journal of Behavioral Development 25 (2) (2001) 140–147.

[8] D. Zarrella, Is 22 Tweets-Per-Day the Optimum? (2009). URL `https://blog.hubspot.com/blog/tabid/6307/bid/4594/Is-22-Tweets-Per-Day-the-Optimum.aspx`

[9] K. Kokalitcheva, WhatsApp now has 700M users, sending 30B messages per day (2015). URL `http://venturebeat.com/2015/01/06/whatsapp-now-has-700m-users-sending-30b-messages-per-day/`

[10] C. Wilke, S. Richly, S. Götz, C. Piechnick, U. Aßmann, Energy Consumption and Efficiency in Mobile Applications: A User Feedback Study, in: Proceedings of The 2013 IEEE International Conference on Green Computing and Communications (GreenCom 2013), IEEE Computer Society, 2013, pp. 134–141.

[11] J. Schalkwyk, D. Beeferman, F. Beaufays, B. Byrne, C. Chelba, M. Cohen, M. Kamvar, B. Strope, Your Word is my Command: Google Search by Voice: A Case Study, in: Advances in Speech Recognition, Springer, 2010, pp. 61–90.

[12] J. Aron, How innovative is Apple's new voice assistant, Siri?, New Scientist 212 (2836) (2011) 24.

[13] R. Trestian, A. A.-N. Moldovan, O. Ormond, G. Muntean, Energy consumption analysis of video streaming to android mobile devices, in: Network Operations and Management Symposium (NOMS), 2012 IEEE, IEEE, 2012, pp. 444–452.

[14] N. Thiagarajan, G. Aggarwal, A. Nicoara, D. Boneh, J. P. Singh, Who killed my battery?: analyzing mobile browser energy consumption, in: Proceedings of the 21st international conference on World Wide Web, ACM, Lyon, France, 2012, pp. 41–50.

[15] G. Kalic, I. Bojic, M. Kusek, Energy consumption in android phones when using wireless communication technologies, in: MIPRO, 2012 Proceedings of the 35th International Convention, IEEE, 2012, pp. 754–759.

[16] E. J. Vergara, S. Andersson, S. Nadjm-Tehrani, When Mice Consume Like Elephants: Instant Messaging Applications, in: Proceedings of the 5th international conference on Future energy systems (e-Energy'14), Vol. 1, 2014, pp. 97–107.

[17] J. Newman, Peak Battery: Why Smartphone Battery Life Still Stinks, and Will for Years.

[18] A. Alexiou, M. Haardt, Smart antenna technologies for future wireless systems: trends and challenges, Communications Magazine, IEEE (September) (2004) 90–97.

[19] G. Hanington, P.-F. Chen, P. Asbeck, L. Larson, High-efficiency power amplifier using dynamic power-supply voltage for cdma applications, Microwave Theory and Techniques, IEEE Transactions on 47 (8) (1999) 1471–1476. doi:10.1109/22.780397.

[20] S. Stapleton, F. Costescu, An adaptive predistorter for a power amplifier based on adjacent channel emissions [mobile communications], Vehicular Technology, IEEE Transactions on 41 (1) (1992) 49–56. doi:10.1109/25.120144.

[21] T. Page, Usability of text input interfaces in smartphones, Journal of Design Research 11 (1) (2013) 39–56.

[22] A. Carroll, G. Heiser, An Analysis of Power Consumption in a Smartphone, USENIX annual technical conference (2010) 271–285.

[23] C. Yoon, D. Kim, W. Jung, C. Kang, H. Cha, Appscope: Application energy metering framework for android smartphone using kernel activity monitoring., in: USENIX Annual Technical Conference, 2012, pp. 387–400.
[24] G. P. Perrucci, F. H. P. Fitzek, J. Widmer, Survey on energy consumption entities on the smartphone platform, in: Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd, IEEE, 2011, pp. 1–6.
[25] E. J. Vergara, S. Nadjm-Tehrani, M. Prihodko, Energybox: Disclosing the wireless transmission energy cost for mobile devices, Sustainable Computing: Informatics and Systems 4 (2) (2014) 118 – 135, special Issue on Selected papers from EE-LSDS2013 Conference. doi:http://dx.doi.org/10.1016/j.suscom.2014.03.008.
[26] F. Jiang, E. Zarepour, M. Hassan, A. Seneviratne, P. Mohapatra, When to type, talk, or swype: Characterizing energy consumption of mobile input modalities, in: Pervasive Computing and Communications (PerCom), 2015 IEEE International Conference on, 2015, pp. 114–122. doi:10.1109/PERCOM.2015.7146517.
[27] Android, Android sdk reference of batterymanager2014.
URL `http://developer.android.com/reference/android/os/BatteryManager.html`
[28] Currentwidget : An android widget that display the electric current usage of the device, Online.
URL `http://code.google.com/p/currentwidget`
[29] E. Cuervo, A. Balasubramanian, D.-k. Cho, A. Wolman, S. Saroiu, R. Chandra, P. Bahl, MAUI: making smartphones last longer with code offload, in: Proceedings of the 8th International Conference on Mobile systems, applications, and services, MobiSys '10, New York, USA, 2010, pp. 49–62.
[30] B. Zhao, B. C. Tak, G. Cao, Reducing the Delay and Power Consumption of Web Browsing on Smartphones in 3G Networks, in: Proceedings of the 31st International Conference on Distributed Computing Systems (ICDCS), Minneapolis, USA, 2011, pp. 413–422. doi:10.1109/ICDCS.2011.54.
[31] F. Qian, Z. Wang, A. Gerber, Z. M. Mao, S. Sen, O. Spatscheck, Characterizing radio resource allocation for 3g networks, in: Proceedings of the 10th ACM SIGCOMM conference on Internet measurement, ACM, 2010, pp. 137–150.
[32] J. Huang, F. Quian, A. Gerber, Z. M. Mao, S. Sen, O. Spatscheck, A Close Examination of Performance and Power Characteristics of 4G LTE Networks, MobiSys (2012) 225–238.
[33] R. W. Soukoreff, I. S. MacKenzie, Theoretical upper and lower bounds on typing speed using a stylus and soft keyboard, Behaviour & Information Technology (14) (1995) 370–379.
[34] Nielsen, Nielsens Mobile NetView, Tech. rep. (2013).
[35] M. Szell, S. Grauwin, C. Ratti, Contraction of online response to major events, PloS one 9 (2) (2014) e89052.
[36] C.-F. Chiasserini, R. R. Rao, Pulsed battery discharge in communication devices, in: Proceedings of the 5th annual ACM/IEEE international conference on Mobile computing and networking, ACM, 1999, pp. 88–95.
[37] J. O. Wobbrock, B. A. Myers, Analyzing the input stream for character-level errors in unconstrained text entry evaluations, ACM Transactions on Computer-Human Interaction (TOCHI) 13 (4) (2006) 458–489.
[38] D. Klein, J. Smarr, H. Nguyen, C. D. Manning, Named entity recognition with character-level models, in: Proceedings of the seventh conference on Natural language learning at HLT-NAACL 2003-Volume 4, Association for Computational Linguistics, 2003, pp. 180–183.
[39] S. Kaki, E. Sumita, H. Iida, A method for correcting errors in speech recognition using the statistical features of character co-occurrence, in: Proceedings of the 17th international conference on Computational linguistics-Volume 1, Association for Computational Linguistics, 1998, pp. 653–657.