Investigating the Value of Privacy within the Internet of Things

Alex Mayle*, Neda Hajiakhoond Bidoki†, Sina Masnadi†, Ladislau Bölöni† and Damla Turgut†

*Department of Computer Science, Ohio University
Email: am218112@ohio.edu
†Department of Computer Science, University of Central Florida
Email: {hajiakhoond, sina}@knights.ucf.edu, {boloni, turgut}@cs.ucf.edu

Abstract—Many companies within the Internet of Things (IoT) sector rely on the personal data of users to deliver and monetize their services, creating a high demand for personal information. A user can be seen as making a series of transactions, each involving the exchange of personal data for a service. In this paper, we argue that privacy can be described quantitatively, using the game-theoretic concept of value of information (VoI), enabling us to assess whether each exchange is an advantageous one for the user. We introduce PrivacyGate, an extension to the Android operating system built for the purpose of studying privacy of IoT transactions. An example study, and its initial results, are provided to illustrate its capabilities.

I. INTRODUCTION

The paradigm known as the Internet of Things (IoT) is a pervasive one. The technology once embodied by commercial RFID and wireless sensor networks has found its way into consumer electronics, appliances, and homes. Previously inert devices, such as thermostats, watches, refrigerators, and speakers, are now available and becoming more common in people’s everyday lives.

Consequently, it is easier than ever before to collect vast amounts of data about the end-users of these devices. Many IoT service providers leverage this capability, collecting personal information with the goal of delivering or monetizing their product. However, as service providers’ capacity to collect this data increases, so too, does the need to protect user privacy.

The term “privacy” is widely used, yet bewilderment persists over the meaning, value, and scope of the concept [1]. Privacy is a personal matter, and often means different things to different people. For the purposes of this study, we employ Westin’s interpretation [2]: privacy is one’s right to select what personal information is disclosed to others. While this definition is certainly intuitive, it does not address how information is collected using modern IoT technology. For this, we prefer the following definition inspired by Solove [3]: privacy is an individual’s awareness of, and ability to alter, stop, or continue the collection and processing of personal information to any organization or other individual.

With these definitions in mind, it is clear that in order to maintain his or her privacy, a user must play an active role and make responsible decisions. However, in the context of humans interacting with IoT devices, there are many factors that hinder the user’s ability to do so. For instance, many times, users are not even aware that their data is being collected. This can make it difficult to assess how giving up personal data affects their privacy, leading to a hindered ability to make appropriate decisions regarding their privacy at a later point. Moreover, if a user is aware of data collection, it may still be unclear what exact rights are being given to a service provider. For example, an application asking for access to one’s text messages or contacts need not make it clear whether they will be able to send text messages to those contacts, or simply read them. Moreover, after granting a service access to a form of personal data, this preference is often saved. Meaning, when the service would like to access that particular form of personal data again, the user is not explicitly alerted, obfuscating what data is actually being transmitted to service providers.

In any case, many services are genuinely ineffective if some amount of personal information is not supplied. For example, a service which tracks how far one walks would not work very well without the user’s location. In this way, there is a “trade” between the user and the service provider; the user release some of their personal information in exchange for a service. For this relationship to persist, it must be a win-win for both the user and the service provider. Accordingly, there must be some framework to measure how favorable a particular “trade” is for a user or company. Turgut and Bölöni [4] describe value of information (VoI) and cost of privacy (CoP) in IoT area.

Usually companies are able to estimate their cost and benefit; however, for reasons already mentioned, users are ordinarily not in such a position. In this paper, we provide users with a new, quantitative framework for analyzing each “transaction” in this continuous trading process. Further, we provide an extended version of the Android operating system, PrivacyGate, which enables users to see what data is being transmitted to service providers through each transaction. A user study is implemented using the software, in which participants are given monetary offers in exchange for chunks of their personal data. We collect only initial results, but we aim to expand the study, uncovering trends in IoT users’ Value of Privacy (VoP). PrivacyGate, along with these results, will provide both a new way of managing and reasoning about one’s own privacy.

This paper is organized as follows. In Section II, the related work in the area of mobile privacy and VoI is provided. In
Section III, we discuss how we apply the value of information in our application scenario and introduce the PrivacyGate on a conceptual basis before describing its implementation. In Section IV, we present our user study and discuss the results in detail. Section V concludes the paper.

II. RELATED WORK

Game-theoretic concepts have been useful when modeling privacy in mobile economy. However, it is mostly applied to network agents. It investigates if an independent decision maker is cooperative, selfish, or malicious (or anything in between) and the relationship between the decision making of agents and network security [5].

Chorppath et al. [6] use such theories to create a quantitative model of users’ anonymity in terms of the granularity with which they disclose their location. Panaousis et al. [7] develop a framework in a similar fashion to study how to best encourage mobile users to provide their location. Unlike our work, however, these studies focus solely on location data and introduce only theoretical frameworks without soliciting empirical data. Our research is inspired to a higher degree by the application of the game-theoretic concept, value of information, utilized in the field of networking. Applications include value of information based scheduling of cloud computing resources [8], scheduling data retrieval from underwater sensor networks [9], and developing routing protocols for intruder tracking networks [10], [11]. Bisdikian et al. [12] describe a similar concept, Quality of Information, to also measure qualitative attributes in networks.

Attempts have been made in the past to quantify the value of ones private information. Cyrcek et al. [13] investigate the value mobile phone users place on their location privacy as it varies across European countries using a survey. In our application, we consider more than user’s location information while using a real application to put the user in a more realistic situation. Hann et al. [14] explore the cost-benefit trade-off of disclosing private information to websites. They investigated individuals’ preferences over websites with differing privacy policies. The results show that cost-benefit trade-offs did not change with personal characteristics including gender, contextual knowledge, individualism, and trust propensity. A number of other studies cover specific use-cases and utilize various quantifying methods. These include privacy in the context of social media [15] [16] and quantifying the VoP using in-person group auctions [17]. Acquisti et al. [18] re-frame the question by asking users what they would pay to prevent the disclosure of their private data, while Pu et al. [19] explore the value people place on their friends’ privacy. Finally, Braunstein et al. [20] determine the value of privacy without explicitly asking users. We provided a situation for user to freely make a decision if he or she has consent to disclose the information or not in addition to the value they put for their private information. Private information in our application is not limited to location data but it covers gallery, camera, friend list, and so on. Additionally, we make the users to think about their privacy. We implicitly, make them aware of what information their applications can access. They will get informed of the cost that they actually are paying in using a specific service. We essentially aim to educate people about value of privacy while asking them for value of their information. This would be beneficial to IoT companies as well since their success stands on the creation of a business model that both customers and providers are realized as beneficial. This is the fact that has been proved experimentally to be true for all technological innovations.

III. APPLYING THE VALUE OF INFORMATION TO PRIVACY

When we view an IoT user as making a series of transactions, we effectively create a “game” of sorts. The user trades private data for various services in an attempt to maximize the utility they receive, while minimizing the value of privacy loss. It is the latter value that we aim to empirically quantify.

While the value of information is defined as the price an optimal player would pay for a piece of information, the participants of the study may not fit this description. Some users do not fully understand what information is being requested of them, or how it will be used by the service provider. Moreover, participants responses will presumably be the product of personality, education, and past experiences, among other factors. For example, someone who has had their privacy breached may vary well be more conservative with the information they disseminate. For these reasons, our results may reflect a common attitude across IoT users, rather than what an optimal player would value their privacy as.

A. PrivacyGate

Most of today’s mobile operating systems aim for a middle ground between usability and privacy control. Specifically, Android prompts the user for consent before allowing an application to access private data, but once consent has been given, it applies this decision to later requests. Meaning if an individual agrees to provide their location to an application, that same app may query this data without explicitly asking in the future. This benefits users who would like to grant an application access to their data for an extended period of time, as they will not have to provide consent more than once. While many individuals may fall into this category, users who desire to control their private data on a more granular level are left with a diminished ability to do so. Their only option would be to manually revoke the application’s access to private data after each period of consent. This, however, is far from an ideal solution, as users must go on frequent, disrupting tangents to adjust their settings. It is worth noting that this problem is not unique to the Android OS. Users of other mobile operating systems, namely iOS, are also affected.

PrivacyGate takes a far more reserved approach to the control of one’s privacy. Instead of applying a user’s consent to subsequent requests for private data, the system prompts for consent before each transaction.
1) Defining Transactions: It is useful to think of IoT users as engaging in a series of transactions, but it is not natural. In the case of discrete forms of private data, such as call logs and text messages, this transactional model is easily applied. However, continuous data, including location, presents a larger challenge.

We must define the bounds of transactions such that they are intuitive to users. Ideally, an individual will consent to providing sensitive data when it is needed, and this privilege will be taken away when the current task is complete. To define transactions that reflect this, we discuss four conditions on which a user should be prompted for consent.

C1 - The app has not been used in X time: As the time between sequential interactions with an application increases, it becomes more likely that the two occurrences are unrelated. In this way, a user can provide private data in one context, and be assured that this privilege is not presumptuously extended during subsequent, disjoint tasks.

C2 - The user has not been prompted for consent in Y time: A user may interact with an application often enough such that the threshold X is never reached. While this prevents individuals from being faced with unnecessary prompts during prolonged tasks, they may still desire a less frequent mechanism to control their privacy. This acts as a safeguard, assuring users’ choice to disclose private data is a conscious one.

C3 - The app was force closed by the user: If a user explicitly force closes an application, it is clear that whatever task they were working on is now over. Because of this, they will need to provide consent at some point after reopening the application.

C4 - The device was restarted: This provides a quite obvious separation between tasks. As such, all applications’ access to private data is revoked, requiring the user to provide consent at some point after the device has been restarted for each app.

X and Y are non negative values such that X is less than Y. The definition of a transaction, and the number of prompts, rely heavily upon these values. Therefore, these may be adjusted to achieve the desired balance between usability and privacy control.

2) PrivacyGate VOI Model: To quantify the cost of disseminating personal information and the value received from the service as a consequence of doing so, we propose two functions, \( VoP(c_n, u) \) and \( VoI(c_n, u) \), respectively. We define \( VoP(c_n, u) \) as follows:

\[
VoP(c_n, u) = \begin{cases} 
  o_{c_n,u}/p_{i,u} & \text{if } ac_{c_n,u} = 1 \\
  \infty & \text{Otherwise} \end{cases} 
\]  

where \( ac_{c_n,u} \) is the amount of money which the applications offers the user during transaction \( c_n \). Concretely, this amount of money is offered to the user exactly when the service requests access to a particular form of data. These transactions can occur while the user is engaged with any kind of service, so long as that service requests access to user data. It is reasonable to assume that each user has a unique set of priorities regarding their data. For example, a user may be more reluctant to disclose their location, but have no problem granting access to their contact list. To be able to compare the amount of money accepted by user, we define a user’s privacy unit cost as \( o_{c_n,u}/p_{i,u} \) in the case where the offer is accepted.

Now, suppose a user is engaged with a service and that service requests access to a form of user data. The user will then be presented with an offer. If the user accepts the money offered, we may then say that the user discloses the privacy for \( o_{c_n,u}/p_{i,u} \). On the other hand, if the user rejects the offer we consider the new privacy cost to be infinity. This case can be interpreted as the user rejecting the money in favor of preserving their privacy at that point in time.

It is important to note that while some users may simply never accept a monetary offer in favor of preserving privacy, there are compelling reasons for a user’s behavior to differ with each transaction. For example, consider the case in which a user is at location \( A \), whereupon they receive an offer and they accept it. Subsequently, they travel to location \( B \) where they receive an offer, but this time they reject it. This may be reasonable if, for example, the first offer is enough money to convince them to disclose their location, but the second offer is much less. Furthermore, the user may not care if people know they are at location \( A \), but would rather people be unaware that they were at location \( B \). This could be because of the nature of the location, or any number of personal reasons. Regarding \( VoP(c_n, u) \), we define \( VoI(c_n, u) \) as follows: \( VoI(c_1, u) = VoP(c_1, u) \)

\[
VoI(c_n, u) = \begin{cases} 
  \min(VoI(c_{n-1}, u), VoP(c_n, u)) & \text{if } ac_{c_n,u} = 1 \\
  \infty & \text{Otherwise} \end{cases} 
\]
During first transaction \((c_1)\), we do not have any prior information regarding the user. We judge solely based on their current action. If they reject the offer, we conclude that they prefer to preserve their corresponding personal information, otherwise they are selling the information for \(o_{c_i,u}/p_{i,u}\) unit of money.

After a number of transactions, we have a history of a user’s actions. Their new action may then give us a new piece of information or not. If the user rejects the offer, again, we conclude that during the current transaction, the particular user data that was requested is more valuable to them than the amount of money offered. Regardless of their previous actions, this information is new for us. We know that the amount of money offered is not enough for them to disclose personal data. Again, they may have rejected for numerous reasons, such as being in a private place. On the other hand if the user accepts the offer, it means that they would sell the requested personal data. The amount of money does not differ highly, if the user had already sold his or her information for a smaller amount of money, it does not give us a new information. If they were already comfortable selling their data for a smaller amount, it follows that they would also do it for a larger amount. However, if they sell it for a smaller amount of money during the current transaction than they had previously, it means they are comfortable selling their information for this lesser amount.

Fig. 1 shows the value of information for two random users. It can be seen that the value of information and the value of privacy for the users can vary significantly over time. It is interesting to inspect if the user cares more about specific privacy items or not. For example, if the user holds a higher priority over keeping their location information than their call logs. We will define a new value of information function which considers the privacy item as an input. The new VoP and VoI definitions are as follows.

\[
\begin{align*}
\text{VoP}_i(c_{i,n}, u) &= \begin{cases} 
 o_{c_{i,n},u} & \text{if } ac_{c_{i,n},u} = 1 \\
 \infty & \text{Otherwise} 
\end{cases} \\
\text{VoI}_i(c_{i,1}, u) &= \text{VoP}_i(c_{i,1}, u) \\
\text{VoI}_i(c_{i,n}, u) &= \begin{cases} 
 \min(\text{VoI}_i(c_{i,n-1}, u), \text{VoP}_i(c_{i,n}, u)) & \text{if } ac_{c_{i,n},u} = 1 \\
 \infty & \text{Otherwise} 
\end{cases}
\end{align*}
\]

Using the new VoP and VoI functions, we are able to observe the user’s privacy cost trend with regards to each type of personal information.

IV. IMPLEMENTATION AND RESULTS

A. Implementation

The Android operating system implements a robust system of permissions, used to control access to private user data. When a permission is given to an application to access some form of information, it is not revoked until the user manually adjusts their settings. PrivacyGate merely implements an alternative scheme for automatically revoking permissions.

If an application does not currently have permission to a source of user data, a dialog will appear to request the user’s consent. However, if the application holds the correct permission, the user will have no way of knowing that their data is being accessed. Meaning, to enable the control of privacy on a transactional basis, the operating system must be more proactive in revoking applications’ permissions.

To prompt the user for consent in accordance with C1 - C4, hooks tasked with revoking permissions are placed throughout the ActivityManager. This component of the Android Framework is called upon to open, close, and navigate through applications. Whenever a user interface appears, including when the device is unlocked, C1 and C2 are checked. A specific function is called when an application has been force closed, providing a perfect place to check for C3. Finally, C4 is satisfied in a roundabout way: instead of revoking permissions, PrivacyGate prevents the system from saving permission changes to persistent storage. As permissions are denied by default, this guarantees that applications will not be granted access to private data upon restarting a device.

The PackageManager, responsible for managing applications and their permissions, exposes an extended API. The aforementioned hooks utilize this to revoke applications’ permissions.

There does exist one major limitation within the implementation of PrivacyGate. Applications designed for versions of Android earlier than 6.0 do not request permissions at runtime. Instead they must be granted before installation is completed. This means that apps designed for these earlier Android version would simply crash if governed by PrivacyGate. This limitation will become increasingly insignificant as more service providers update their applications to the latest version of Android.

B. Results

To assess the value of privacy in its many forms, a user study is conducted. Users were selected among the graduate students within the Computer Science Department of the University of Central Florida. In the beginning, many students were willing to participate in the survey. This is usually the case when a student needs to conduct a survey for their research; however, when more explanation was provided regarding the nature of the survey and the phone itself, students became reluctant. Ultimately, many students decided that they would rather not participate. This is interesting because the phone is quite literally the same as any Android device, the only difference being that they can choose to control their privacy on a more granular level. Regardless, the study was conducted with the students who felt comfortable participating.

Throughout the survey, instead of applications merely requesting the consent of the user, monetary offers are provided. Participants may decide to consent and receive the reward, granting the application access to their data, or deny the offer and prevent the application from accessing the requested information. These
offers are generated randomly so we may explore the value for which users willingly provide their private data.

The results consist of five participants over the course of a day for each participant. Participants used an LG Nexus 5X, running PrivacyGate, for the duration of the study. The offers ranged from $0.00 to $2.00. The values of $X$ and $Y$ were set to forty and ninety minutes, respectively.

Because it is up to the application to decide when it accesses user information, it also dictates when a prompt for consent will appear. This means that a keen participant may learn what steps to take in an application to produce such a dialog, and the monetary offer that is presented along with it. To stop them from capitalizing on this, a wait period of twelve minutes is enforced if the user declines the offer. This prevents a user, who would have otherwise accepted the offer, from denying it with the intention of generating an offer of higher value.

It is interesting to see how many of offers for each privacy item were accepted by a user. This shows how a user prioritizes the types of personal information they would like to keep secure. Moreover, it illustrates whether a user indiscriminately accepts or denies offers of a particular privacy item or shows varying behavior throughout the day. Table II illustrates acceptance rates for five different users who participated in our study, classified by permission type.

Considering location, User 1 was very conservative, denying the request every time, in turn refusing to disseminate the location to the app. In contrast, User 3 and User 5 always accepted the offer, receiving the reward and releasing the requested data. Somewhere in between these two poles, we have User 2 and User 4, who accepted 25% and 60% of the offers, respectively.

As we can see, users exhibit very different behavior regarding their decisions to accept or deny each offer. Aggregating the offers for each privacy item together, User 1 accepted 62.5% of the offers, while User 2 accepted only 34% of them. As the data on Table II suggests, User 3 and User 5 are more similar to User 1, in that they both show high acceptance rate. The difference between User 1 and User 5 is that User 1 is more cautious on the files, and then rejected the next offer that requested the information. User 4 is more reluctant to share the Location and Storage data when compared to the other forms of personal information. We can see that different users exhibit a unique prioritization of their personal information and take the steps needed to protect the data they care about the most.

<table>
<thead>
<tr>
<th>Permission</th>
<th>User 1</th>
<th>User 2</th>
<th>User 3</th>
<th>User 4</th>
<th>User 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td>0%</td>
<td>25%</td>
<td>100%</td>
<td>60%</td>
<td>100%</td>
</tr>
<tr>
<td>Camera</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Storage</td>
<td>50%</td>
<td>75%</td>
<td>10%</td>
<td>66%</td>
<td>30%</td>
</tr>
<tr>
<td>Contacts</td>
<td>100%</td>
<td>17%</td>
<td>66%</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Figure 1 represents two users decision changes in addition to the perceived VoI. We can see that the opinion of each participant has changed over time. In several cases, consecutive offers for the same privacy item solicited different responses for the user. In other words, the user accepted one offer for access to the files, and then rejected the next offer that requested the same information. Corresponding VoIs represents how different the value of privacy changes for each user. Furthermore, when we compare the two users, we can see that their responses to the offers vary greatly. This further supports the notion that each
user has a unique characterization of what personal information is important to them. We are also interested in knowing how a single user responds to requests for varying types of personal data throughout one day. Figure 2 depicts two participants’ decisions regarding access to their device files.

TABLE III
PERMISSION PRIORITY FOR FIVE DIFFERENT USERS IN DESCENDING ORDER.

<table>
<thead>
<tr>
<th>User 1</th>
<th>User 2</th>
<th>User 3</th>
<th>User 4</th>
<th>User 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td>Contacts</td>
<td>Storage</td>
<td>Location</td>
<td>Storage</td>
</tr>
<tr>
<td>Storage</td>
<td>Camera</td>
<td>Contacts</td>
<td>Storage</td>
<td>Camera</td>
</tr>
<tr>
<td>Camera</td>
<td>Location</td>
<td>Location</td>
<td>Camera</td>
<td>Location</td>
</tr>
</tbody>
</table>

Prioritization of different privacy items can also be different for each distinct user. Table III shows participants priorities for each permission in a descending order. As the data suggests, Storage permission is among the top two for all the participants. This means that users are more cautious about letting an app access the files on their device, though each user has a different order of priority for privacy items.

As we saw from the results, even with a small sample of population, it is clear that different users can have totally different action regarding their personal information disclosure. They have different actions as time goes, even in a short period of time such as a single day. However, what happens while using the phones in the market is that, the application asks for user privacy item access permission and keeps it for a long time, usually as long as the application is installed on the device. Our results warns that this may not be a privacy respectful fact that exists. It is highly recommended to researchers as well as providers to look for a more appropriate techniques which can guarantee mutual benefits for the users and the providers.

V. CONCLUSION

This paper investigates the value of IoT users’ privacy by modeling their behavior as a series of transactions. Each of which involves the exchange of private data for some service. We introduce PrivacyGate, a mobile OS enabling users to control their privacy on such basis. Using this, and the value of information, we conduct user studies to quantify the value of privacy. Our result showed that the value of privacy of an individual changes as the time moves on. It is also different for different users. We conclude that taking one’s permission to access his personal information and then keeping it for a long period of time as is done by many devices may not be a good idea. Our aim is to not only provide a practical means, but also a theoretical framework for the responsible management of one’s privacy.

We intend to expand the user study such that we may evaluate VoP in its many forms. A participant’s VoP data may be dictated by factors such as age, sex, background, and education. Therefore, it is important that our sample size is sufficiently large and diverse, allowing us to generalize to the greater population of IoT users. In addition to quantifying participants’ VoP, we may also ascertain the value with which they regard their services. While this is highly dependent on the context in which they use them, this information may be applied to the evaluation of each transaction. Through this, we will provide IoT users with a new paradigm for privacy decision making.

Acknowledgements: The support for this work was provided by the National Science Foundation REU program under Award No. 1566302. Any opinions, findings, and conclusions and recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

REFERENCES