RUHR-UNIVERSITÄT BOCHUM

Horst Görtz Institute for IT Security

Technical Report TR-HGI-2013-002

On the Usability of Secure GUIs

Atanas Filyanov, Aysegül Nas, Melanie Volkamer, Marcel Winandy

---

HGI System Security Lab
Ruhr-University Bochum

RUHR
UNIVERSITÄT
BOCHUM

RUB

Ruhr-Universität Bochum
Horst Görtz Institute for IT Security
D-44780 Bochum, Germany

HGI-TR-2013-002
30.09.2013

On the Usability of Secure GUIs

Atanas Filyanov
HGI System Security Lab
Ruhr-University Bochum
Germany
atanas.filyanov@trust.rub.de

Aysegül Nas
HGI System Security Lab
Ruhr-University Bochum
Germany
ayseguel.nas@trust.rub.de

Melanie Volkamer
Dpmt of Computer Science
TU Darmstadt
Germany
melanie.volkamer@cased.de

Marcel Winandy
HGI System Security Lab
Ruhr-University Bochum
Germany
marcel.winandy@trust.rub.de

ABSTRACT

Secure GUIs have been proposed in the literature and there are already a few operating systems which enable secure GUIs. The main idea is that a trusted part of the operating system controls what is displayed on the screen. Most of the secure GUI proposals include a reserved area on the screen that is used to display information about which application is currently having the input/output focus of the user and what type of trustworthiness this application has. Obviously, it is very important that users know the meaning of this reserved area and only edit or enter sensitive data if the application with the corresponding label is having the focus. However, whether this assumptions holds for average users has not been evaluated to the best of our knowledge. With our research we try to shed light in this situation. We evaluated in a lab user study two different approaches to display the reserved area as trusted statusbar. Our results show that the trusted statusbar—independent from being displayed on the top or the bottom of the screen—enables participants to select the proper application as long as no fake but authentic looking client or software is executed in an untrustworthy application.

Keywords

Secure GUI, Trusted Path, Usable Security, Usability, User Study

1. INTRODUCTION

We use commodity computing platforms for many tasks, including entering or editing sensitive data on them. Unfortunately, the graphical user interfaces (GUI) running on these devices are not designed to provide a secure means of ensuring users that they are interacting with the authentic application and not with some fake one. These design weaknesses are often exploited by adversaries. They try to steal security or privacy sensitive data by tricking users to enter such data into "authentic"-looking applications.

Secure GUIs have been proposed as a solution to this problem, e.g., [5, 6, 7, 9, 10, 17], and few commercial operating systems implement SecureGUIs, e.g., [8, 12]. The main idea is that a trusted part of the operating system controls what is displayed on the screen, and the user is always able to invoke a trusted path to this part.

Most of the secure GUI proposals include a reserved area on the screen that is used to display information about which application is currently having the input/output focus of the user and what type of security or trustworthiness this application has, i.e., the *labeling* of the application (e.g., trusted/untrusted or confidential/secret/topsecret). Essentially, this label information in the reserved area of the Secure GUI is the only trusted path indicator to the users that they are interacting with the authentic application that belongs to this label. The reserved area for displaying label information in the proposed Secure GUIs is usually a top- or bottom-screen (trusted) status bar, e.g., [6, 7, 8, 9, 12, 17].

Previous works have shown that it is technically possible to construct secure operating systems with secure GUIs. For example, consider a system that is supposed to run trusted and untrusted applications. Existing approaches, e.g., [11] can provide isolated execution environments for these two types of applications. It is also possible to ensure or at least verify the integrity of the operating system and trusted applications [1, 15]). Consequently such a system can implement a Secure GUI that always shows in a reserved area the label (trusted/untrusted) of the application having the focus. If the user enters sensitive data only in the application labeled as trusted, unauthorized eavesdropping or manipulation from untrusted applications can be prevented.

Obviously, it is very important that users know the meaning of the reserved area of Secure GUIs and the different labels with different information in this reserved area. They also need to be able to perceive which application currently has the focus, and they should only edit or enter sensitive data if the application with the corresponding label is having the focus.

A potential drawback of existing Secure GUI proposals is that the security indicators are passive ones—i.e. the system does not actively prevent (because it is technically not possible to do so) the user from entering sensitive data in any untrusted potentially malicious application—and it is already known from previous research in usable security [3, 16, 19, 20] that passive indicators do not provide effective protection against attacks in the web browser context.

Correspondingly, we have the following situation in the Secure GUI area: Existing Secure GUI proposals provide strong security guarantees from a technical point of view. However, none of them has been evaluated with respect to

the effective protection for the average user. Thus, to the best of our knowledge, it is not known whether any of them provides an *effective* protection in real world situations.

With our research we try to shed light in this situation. We evaluate in a lab user study two different approaches to display the reserved area as trusted statusbar: one on the top of the screen and one on the bottom, as these are the most common places for status bars in the default settings of most of the operating systems. Our results show that the trusted statusbar—independent from being displayed on the top or the bottom of the screen—enables participants to select the proper application as long as no fake but authentic looking client or software is executed in an untrustworthy application.

The paper is structured as follows. In Section 2, we discuss related work in the area of Secure GUI approaches and related user studies. Afterwards, we briefly introduce the relevant facts for the user study resulting from our research project in Section 3 and explain how we developed the two evaluated trusted statusbar approaches throughout the first phase of this project. In Section 4 we propose the study design, and in Section 5 we present the results of the study. In the last Section, we discuss the results and present plans for future work.

2. RELATED WORK

In this section we outline existing concepts of Secure GUIs to provide a trusted path to the user, summarize related usability studies from usability of web browsers' security indicators, and finally provide an overview or our research project and the architecture of the evaluated system.

2.1 Secure GUI Approaches

Several approaches for realizing a trusted path from the operating system to the user have been proposed over the years, though none enjoy widespread adoption on commodity systems. We briefly review some representative designs.

One approach is to use a trusted window manager, as proposed by Epstein [5, 7, 6], where each window is visually labeled and a dedicated area of the screen is reserved for the exclusive use of a trusted software component that shows the label (identity and status) of the current application. As this is mainly related to a graphical user interface (GUI), the concept is also referred to as "Secure GUI". We briefly review some implementations of Secure GUI systems:

- **Colored window labels and reserved area**: TX [7, 6] is a multi-level secure X window system. It multiplexes the windows of all levels and attaches a colored label on each window to indicate its security level. A reserved area on the screen always shows the level having the input focus. Trusted Solaris [8] is a commercially available implementation that follows the concepts of Trusted X, whereas EWS [17] and Nitpicker [9] are research prototypes that have different internal implementations, but on the GUI side follow the same principles (window border coloring and reserved area for a trusted status bar).

- **Reserved area status bar only**: Green Hills' INTEGRITY [12] is a microkernel-based operating systems that supports multi-level security. It displays the maximum security level and the current input security level at the top and, respectively, bottom of the screen as a colored bar. Applications in the system are virtual machines (VMs) that run isolated from each other. The VMs can only access virtual devices and cannot draw on the reserved screen areas.

- **Split screen**: The SDH architecture [18] divides the screen in fixed separate regions according to security levels. Applications of the same level are always shown only in the corresponding fixed region on the screen. While this approach may be able to avoid confusion about where to enter sensitive data, it clearly limits the screen size available to the applications.

While the Secure GUI concept has been implemented in some (mainly research) systems, the goal of widespread adoption has still remained elusive.

A closely related variant leverages the notion of a **secure attention sequence**, e.g., "Press Control-Alt-Delete to log on." The assumption here is that the OS kernel remains uncompromised, and will always be the first software layer to process keyboard input. Thus, any spoofed login dialog box will be immediately overwritten by the legitimate box. However, users must be taught to always press the necessary key sequence, and this functionality is limited to operating system specific actions.

A final approach is to use some form of **dedicated additional hardware** as an axiomatically trustworthy indicator, in the limit something as simple as a dual-color LED [14, 13]. This design is compelling as it still enables full screen applications, which must otherwise be disallowed given their ability to spoof other security indicators.

2.2 Related Usability Studies

We studied existing literature on passive warnings in order to compare their applicability to Secure GUIs. Essentially, existing studies mainly concentrate on web browser security and target phishing attacks.

Several studies already exist in the literature (examples are: [3, 16, 19, 20]) that evaluate security indicators (such as the URL bar and the lock icon) of existing web browsers: Dhamija et. al. analyzed what makes a phishing web site look trustworthy [3]. They have found that even experienced users can be tricked by easily modified visual objects. In our study, we want to see how do security indicators help users determine if they are working in a trustworthy compartment and if users actually look for such indicators.

The usability study presented in [16] shows that users do not pay attention if *https* is missing, and enter their passwords. Moreover, they discard the security hints presented by the browser, e.g. that the web site certificate is wrong, and give away their passwords again.

In [19] they performed a user study, aiming to test which security indicators in browsers (in their case Internet Explorer) are perceived by the users, which are ignored, and if it is easy for users to find these indicators or it is rather difficult. Such results were determined using eye trackers and later on user surveys. The end results showed that the most useful symbol is the *lock* symbol, showing when connection is secured. However, only few users interacted with this symbol. Moreover, information about the web site certificate is very rarely reviewed. Further, users stopped looking for security indications once they were logged in.

Another study [20] tested few different security toolbars. The results of the study showed that such toolbars are rather

ineffective during phishing attacks. It proves the point that security is not the primary goal for the users, and available security indications are not always checked.

In summary, those existing user studies show that passive security indicators apparently are not an effective measure to prevent users from entering sensitive data into untrusted applications. However, the *applications* in those studies are actually *web sites*, or more precisely, web content rendered and running in one and the same application, i.e., the web browser. Our scenario is different, though: We have different actual applications. We want to find out whether passive security indicators can still have a meaning in the context of Secure GUIs, where we need to indicate the *trustworthiness of complete applications* in contrast to the content rendered in just one application.

Closest to our study might be the study in [2], which experiments with different background color lights of the keyboard. Similar to the color of a label in a SecureGUI, the background color of the keyboard can indicate different risk or trustworthiness levels. The results of that study showed that users behaved more secure with this special keyboards. However, this is complementary to our approach as they actually analyze a dedicated extra hardware device.

3. OUR RESEARCH PROJECT

We started our research project with the goal to develop and evaluate a secure operating system for information flow control and protection of sensitive data from unauthorized disclosure and manipulation. The processing of sensitive data should be securely separated from other workflows. We aim to test the system in a real-life setting and, thus, as one example we choose as target an electronic student data administration system that is used at a university.

The data stored in the student administration system does not only include personal information such as the names and addresses of students, but also results of exams, courses taken by the students and other education-related data. Students can access the administration system via a client application that connects to the local network of the university. For authentication purposes the students have to provide a smartcard to the client computer and enter their corresponding secret PIN into the client application. Hence, our goal is to protect the process of entering the PIN and any operation on the student data from other (potentially malicious) applications on the client computer. Therefore, our operating system provides separated execution environments for (a) the university client application that is allowed to access the university server, and (b) other applications that are allowed to access the Internet. To enable the users to distinguish between these different applications (and from preventing them to enter their PIN into the wrong application), our operating system implements a Secure GUI.

For a high level architecture of our tested system, see Figure 1. It is based on our prior development during other projects [1]. The main idea is that computer applications run in different *compartments*. The compartments are technically virtual machines (VMs), running a commodity operating system and operating as isolated execution environments on the same platform. Each compartment is assigned a label, i.e., a name and a color. The label serves as visual

[1] *We will add references in the final version – blinded for submission here*

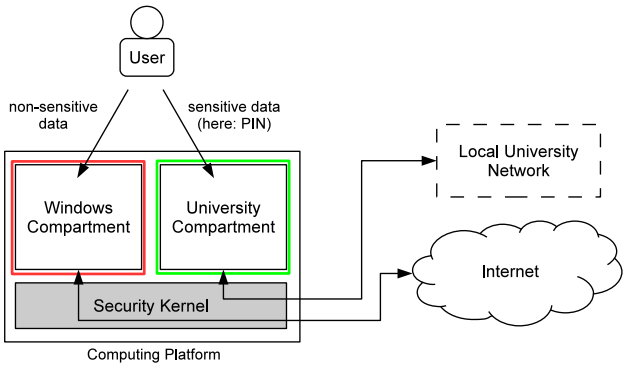

Figure 1: System architecture of our prototype

indication for the authenticity of each compartment.

Our operating system follows the role of a security kernel that provides the following functionality. It:

- runs virtual machines as compartments, only one compartment is always shown in fullscreen, i.e., the one that has the input/output focus; [2]
- controls which compartment is able to access which network sites;
- checks the integrity of the university compartment at startup to ensure its authenticity; and
- provides a Secure GUI implementation with a reserved area ("TrustBar") to indicate the label (name, color, and security information) of the currently active compartment (i.e., the one having the focus).

3.1 Project Settings

In the first phase of our case study we prepared 130 laptops. We preinstalled our tested system and four compartments (WorkWindows, WorkLinux, Students, and Internet). In that way, our participants have the possibility to choose between Windows and Linux for their private working environment (or use both simultaneously). A dedicated compartment allows to access the student administration system in the local university network. It is subject to an enforced security policy, which restricts the incoming and outgoing network connection, allowing only servers of the university to be reached. Note that the screenshot in Figure 2 shows the version that the students received. For our test we used a slightly modified version. More details are given in Section 4.

3.2 Potential Study Participants

To recruit interested students we distributed flyers around the campus and published a press report. As a condition we stated that if selected, interested students need to actively participate in online questionnaires and case studies. In return, these students could use the given laptop after the project end until the end of their studies. Until the specified deadline, 414 students from the university have registered

[2] There is a compartment overview screen showing all available and running compartments, see Figure 2.

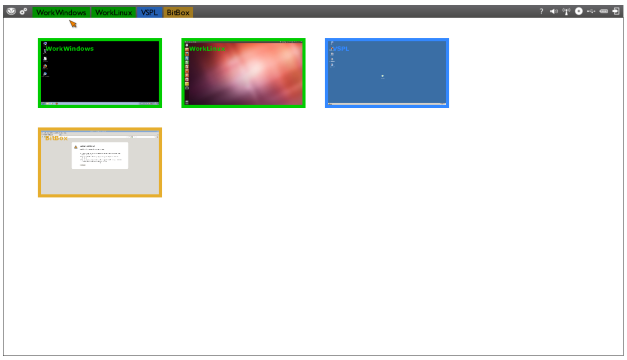

Figure 2: Overview screen of our prototype

their interest in our project. Among those student, we have selected 130 who could participate in our case study. We aimed to have such a student distribution, that all faculties are represented. As a result our participants are enrolled in 36 different study programs. Their ages vary between 18 and 35 (average is 21), with 50% males and 50% females.

3.3 Introductory Course for the Participants

To inform the participants about our project and to distribute the test laptops, we organized an introductory course. Around 120 students took part in it. During the course we gave only a high level information about the tested system. We told the students that there are different compartments installed and one is particularly secure for accessing the electronic student administration system.

We gave them a few hints with respect to working with the operating system as well. We published the given presentation slides on our local university web site, so that they could take a look again later on if they wanted to. We stated that our tested system was still in development and some functional features were missing, and we would continuously deliver system updates and other improvements throughout the project based on the input that we receive from the students.

At the end of the presentation, each of our participants received one of the laptops. Together with the university's privacy officer, we had constructed a relevant consent form, that stated that the participation is voluntary, and we asked every participant to sign it. It was stated specifically which data will be collected and which activity will be monitored and that every private data will be appropriately anonymized and encrypted.

3.4 Considerations for Usable Secure GUIs

Based on results from existing studies in usable security, existing Secure GUI implementations, and a questionnaire and pretests we developed two options for our Secure GUI: one showing a reserved area TrustBar at the top of the screen, and one showing the TrustBar at the bottom of the screen. 101 students filled out the questionnaire and 24 people participated in the pretests.

Moreover, our analysis of the questionnaire showed that the majority of participants consider traffic light colors to best represent compartments with different security properties—i.e., in our case green for the university compartment and red for the normal Windows compartment.

Finally, we decided to add a security hint into the compartment label, next to the compartment name, to make it obvious what is allowed to do in that particular compartment and what is not recommended: "No PIN Entry!" for the Windows compartment, and "PIN Entry secured!" for the university compartment—reflecting the objective to enter the PIN only in the trusted university compartment. See also Figure 4 and Figure 5 for screenshots.

4. STUDY DESIGN

In this section, we describe the research question and study instruments, methodology for data analysis, participant recruitment and group assignment, as well as ethical considerations for the study.

4.1 Research Question

The interfaces proposed in subsection 3.4 were evaluated throughout a lab user study. In particular the following *research question* should be answered: Are students in our project more likely to select the proper compartment with the TrustBar on the top of the screen or with the TrustBar on the bottom of the screen. As we only consider two compartments for the user study, we define proper selection in the following way: Students select the proper compartment if they enter their student credentials (namely student ID and corresponding PIN) only in the university compartment and use the Windows compartment for other tasks, in particular those for which they need to connect to arbitrary Internet servers.

We study three different cases:

- **(case - easy, privacy requiring situation)** Participants are in the Windows compartment and in order to be able to conduct the next task they have to switch first to the university compartment before entering their student credentials. Note, the authentic university client application cannot be started in the Windows compartment. Participants would see an error message in the web browser if they try to do so.
- **(case - flexibility requiring situation)** Participants are in the university compartment and in order to be able to conduct the next task they have to switch first to the Windows compartment. Due to the limited functionality in the university compartment, it is not possible to successfully conduct the task in this compartment. Participants would see an error message in the web browser if they try to do so.
- **(case - difficult, privacy requiring situation)** Participants are in the Windows compartment and in order to be able to conduct the next task they have to switch first to the university compartment before entering their student credentials. However, this time a fake university client application is provided in the Windows compartment which makes it more challenging for participants to switch to the university compartment as it looks like they can also conduct the task successfully in the Windows compartment.

4.2 Study Settings

We evaluated our research question within a lab study using a test laptop with the new interfaces. The participants

worked with either of the two new interfaces. They were told that they cannot use their own laptops because we wanted to test possible future interfaces which are not yet deployed. However, they had to use their own student IDs, student ID cards and PINs to participate. The study consists of the **following parts**:

- The students were welcomed and informed about the purpose of this study. The supervisor reminded the participants about the signed consent from the beginning of the project.
- The students received two pages of instructions describing and explaining the overall scenario and the corresponding tasks to be conducted.
- The students solved a number of tasks on the test laptop. After every task, each participant evaluated the (completed) task, by answering questions on a second laptop. Note, the reason for using a second laptop was to have time to adjust the test laptop unnoticed between two tasks.
- The students filled out a questionnaire about their demographics and they were asked to not reveal any information about the study.

Note, we continuously checked all the available communication channels which have been set up for the project to be sure that the students did not reveal any relevant information about the study over these channels.

Test laptop. For the user study we installed the system with one of the two evaluated Secure GUI interfaces. This laptop had a smart card reader attached. Furthermore, two compartments were already started - namely university compartment and Windows compartment - the compartment overview screen (see Figure 3) was initially displayed at the beginning of the test. Note, we decided to stick to only these two compartments for our test. This setting allowed us to limit the number of possible compartments to conduct the different tasks in the user study. We expected that participants would stick to those compartments which have already been started, which turned to be true during our test. None of our participants in the user study started any of the other available compartments.

In order to evaluate the third case for our research question, we developed a fake university client application, which we installed in the Windows compartment. It is a JavaScript `.hta` application, which looks exactly the same as the original university client application when it displays the key pad for entering the PIN to get access to any further information. In case the participants enter their PIN in this fake university client application and press the 'Send" button on the screen or the "Enter" key on the keyboard, the application displays an error dialog window saying "Connection Error" but does neither store nor send the PIN or any information. Thus, if students entered their PIN here because they did not notice that they have been in the wrong compartment, nothing bad would happen to them and their sensitive data.

Scenario and Tasks. The participants were asked to conduct four different tasks. The different tasks were all related to preparing for a student job application. This is a realistic scenario for students as many students are working part-time and are searching for new jobs from time to time to broader their knowledge and experiences. As many employers ask for a transcript of record and a current matriculation certificate (to check whether they are still students), we included downloading these two documents in our scenario. Note, for both documents, students need to use the university compartment and they first had to login with their student ID card and their PIN. Note, entering the PIN in a fake university client application would mean that the attacker got access to the PIN of the student card. Besides these two critical tasks, we added two less critical tasks. One that can be performed in any of the two compartments and one that can only be performed in the Windows compartment. In detail, we asked the participants to conduct the following four tasks:

Task 1 - Search in local net: Motivated by the fact that speaking foreign languages becomes more and more important and the participants having noticed to have better chances in future job application, the participants were asked to check whether adequate language course are offered at the university in the next semester.
Proper Execution: As the corresponding page is provided by the university both compartments can be used to properly execute this task.

Task 2 - PIN entry w/o attack: Afterwards, the participant were asked to download a recent transcript of records. In order to conduct this task they had to login at the university client application.
Proper Execution: The participants use the university compartment in order to conduct this task. If they were in the Windows compartment after the first task, they would switch to the university compartment without first having tried to use the faked university client application in the Windows compartment.

Task 3 - Search in Internet: Next, we asked the participants to search for a photo studio near their home address in order to be able to include a professional photo in their CV.
Proper Execution: The participants use the Windows compartment in order to conduct this task. If they were in the university compartment after the second task, they would switch to the Windows compartment without first having tried to use a search engine on the Internet in the university compartment.

Task 4 - PIN entry w/ attack In task 4, we asked the participants to download their matriculation certificate to include it in the application and to finalize their job application. In order to conduct this task they had to login at the university client application. Note, during the time the participant answered the questions to task 3 on the second laptop, we launched unnoticed the fake university client application in the Windows compartment in case the participant finished task 3 in that compartment. When the participants turned back to the test laptop, the fake university client application was already running (see Figure 5). If the participants finished task 3 in the university compartment we would not modify the test laptop.
Proper Execution: The participants use the university compartment. If they were in the Windows compartment after the first task, they would switch to the uni-

versity compartment without first having tried to use the university client application or even trying to enter the PIN in the fake university client application in the Windows compartment.

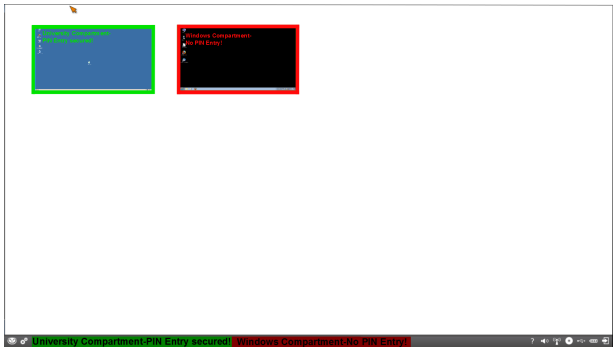

Figure 3: Compartment overview screen

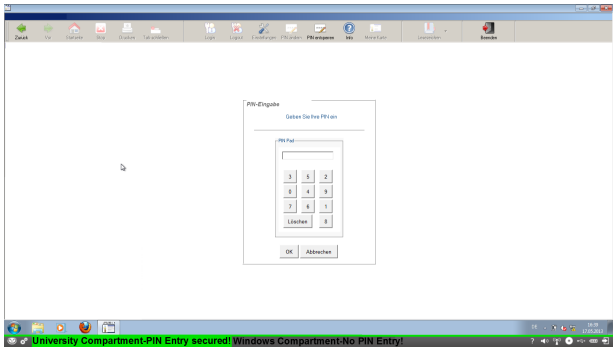

Figure 4: University compartment is active (authentic university client application).

4.3 Data Collection and Analysis

The participants had to answer a few questions after each task and at the end of the study. These questions were concerning the scenario in general and the particular task just conducted. These questions were integrated to convince the participants after each task to go to the second laptop, and thereby enabling the study conductor to start the fake university client application after task 3 without being noticed by the participants. Correspondingly, these questions are not evaluated. The demographic questionnaire included questions on age, gender and subject as well as how often they use the provided laptop, whether they joined the introductory course, and where they are used to see the toolbar. In addition, at the end of the user study, they were asked to explain for each task why they performed each tasks in the way they performed the task.

We also asked participants to think aloud during task processing. The study conductor took notes about successful and failed tasks (without the participants having access to the notes during the study) based on the participants' comments and observing the monitor while the participants conduct the tasks. The study conductor also noted if participants tried to conduct the task in one compartment, failed,

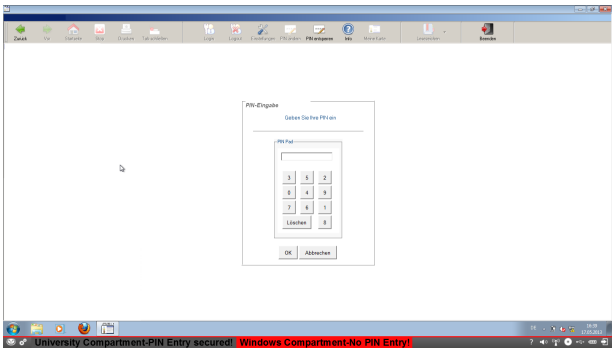

Figure 5: Windows compartment is active (fake university client application).

noticed the problem and then switched to the other compartment to complete the tasks there.

In order to **analyze** the three parts of the research question (including to evaluate whether one of the two approaches performs better) we analyse the following data for both groups:

- number of participants who decided to switch to the proper compartment for each of the tasks;[3]
- number of participants who always conducted all tasks properly (i.e. they always selected the proper compartment); and
- number of participants who decided to select Windows compartment for the first task and number of participants who decided to select university compartment for the first task.

Note, we decided to not count the total number of participants who used the proper compartment for each task because some were already in the proper compartment before. This happened if they selected the wrong compartment for the previous task and thus failed in conducting the previous tasks.

While we do not analyse task 1 (Search in local net) to answer our research question we included this task to see which compartment participants go more likely for if both are possible: the more secure university compartment or the more flexible Windows compartment providing more functionality.

4.4 Participant Recruitment and Group Assignment

All of the 130 participants in the project were invited by email to register for the conducted lab user study. From the registered participants we selected 26. These 26 were assigned to one of the two groups while trying to make sure that the participants in each group represent the sample of the 130 total participants with respect to age, gender, and subject.

According to Dumas and Redish [4], in order to have meaningful results, we would need at least six people. Even, we are aware that more participants would provide more reliable results we agreed to have 13 students in each group

[3]Note, we do not include task 1 in the evaluation as any compartment would have been possible.

due to the early status of this research (comparing two approaches and not a final evaluation of one interface at the end of the project).

The group with the TrustBar displayed at the top of the screen is in the following called *Group-Top* and the one with the TrustBar displayed at the bottom of the screen is called *Group-Bottom*.

4.5 Ethical Considerations

Ethical requirements for research involving human participants are provided by an ethics commission at the university. The relevant ethical requirements (participant consent and data privacy) were met. All students were informed about the purpose of the study and could register for it. However, even they signed at the beginning of the project that they agree to participate in user studies, they could decide not to register for this particular one.

In order to meet the data privacy requirement, a privacy statement was provided on the questionnaire, assuring participants that their data would only be collected for research purposes, their identity would not be linked to their responses, and their data would not be passed on to third parties. Furthermore, participants' data was only handled by researchers involved in the project. This privacy statement for our study participants was confirmed by the data protection officer of the university.

In addition, we do not log the data that participants enter, in particular not the PIN—independent whether they are using the proper compartment (university compartment) or the fake software in the wrong compartment (Windows compartment). While the study conductor took notes about the visited compartments, we did neither observe the keyboard nor the screen while participants entered their PIN.

5. RESULTS

We present the results of our user study in this section.

5.1 Participant Demographics

26 participants (13 female and 13 male) took part in the study. All of them are students at our university. All participants stated that they joined the introductory course and none of them stated that they have problems with color identification.

Group-Top. The Group-Top consists of 13 participants (five males and eight females). Age ranged from 20 to 30 years old. Six study humanities and social sciences, two study natural sciences, two are engineering students and three study medicine. Twelve of the participants normally use the Windows operating system and the other one is Linux user. Twelve Windows users are used to see the toolbar on the bottom of the screen and the Linux user placed the toolbar at the top of the screen. One participant uses our tested system laptop on a daily basis and one uses it several times a week. Two test subjects stated that they use the laptop once a week, three use it several times a month, three use the laptop once a month and three use it only a couple of times a year.

Group-Bottom. From 13 test subjects, eight are males and five are females, with ages between 22 and 36. Eight study humanities and social sciences, three study natural sciences, one is engineering student and one studies medicine. Nine participants are Windows users and four participants use the Linux operating system. All Windows users stated that the toolbar is displayed at the bottom of the screen. Two Linux users positioned their toolbar on the left of the screen and the other two on the top of the screen. Two participants stated that they use the laptop on a daily basis and five use it once a week. Two test subjects use the laptop several times a month, two use it once a month, two use it only a couple of times a year.

Summary. In total, the groups are very similar with respect to all these criteria and represent the 130 project participants very well. All of the test subjects had to state in the final questionnaire how important do they find the PIN. The average stays at 1.96, with 1 being "very important" and 5 being "not important at all". This shows that they should be motivated to select the proper compartment in the user study as they had to use their own student card and corresponding PIN.

5.2 Group-Top

The participants of Group-Top conducted the task on the test laptop displaying the TrustBar at the top of the screen.[4] The different paths participants took throughout the study are shown in Figure 6.

Results relevant for case - easy, privacy requiring situations (Task 2). The number of participants who were in the Windows compartment before task 2 (PIN entry w/o attack) is twelve (see Figure 6). Eleven of them switched to the university compartment to conduct this task (and did not try to access the university client application from Windows compartment). One of twelve participants did not change the compartment and tried to perform this task also there. She entered the PIN code in the fake university client application and received the connection error message. She did not try to execute the task to the end. Seven of eleven participates stated that they switched for security reasons in the Windows compartment and four of them stated that they usually use this compartment for university-related tasks. One participant who was in the university compartment before task 2 conducted also this task in the university compartment. She did so for security reasons.

Results relevant for case - flexibility requiring situations (Task 3). The number of participants who were in the university compartment before task 3 (Search in Internet) is twelve (see Figure 6). Eleven of them switched to the Windows compartment to conduct this task (and did not try to access the Internet/search engine from the university compartment). One of them did not switch to the Windows compartment but she properly decided to switch from Windows compartment to the university compartment in the previous task. She tried to get access to a search engine on the Internet in the university compartment and received the error message "Server not found". She thought there is a problem with the Internet and did not try to exe-

[4]Note that the TrustBar of our underlying security kernel operating system is different from the toolbar that guest operating systems in the virtual machine compartments can display. The TrustBar is always shown in addition and under control of our security kernel.

cute the task to the end. Three of eleven participants stated that they switched for security reasons. Eight of eleven participants stated that they used for such tasks the Windows compartment by habit.

Results relevant for case - difficult, privacy requiring (Task 4). The number of participants who were in the Windows compartment before task 4 (PIN entry w/ attack) is twelve (see Figure 6). Only five of them switched to the university compartment to conduct this task (and did not try to access the university client application from the Windows compartment). The remaining seven participants directly entered their PIN in the started fake university client application. Five of seven participants did not try to execute the task to the end after they received the connection error message. One of the two other participants entered his PIN twice in the started fake university client application, and then he noticed that he was in the Windows compartment. Afterwards he switched to the university compartment and conducted this task there. The other one, after he received the error message, thought that he was in the university compartment. Then he switched to the university compartment and conducted this task there. All five participants who first switched to the university compartment stated at the end of the test that they did that for security reasons. Four of those who directly entered their PIN in the fake university client application stated that they used this compartment accidentally. Two of them stated that they did not noticed they have been in the Windows compartment. It may be that the other participants also have not noticed in which compartment they are located. Three of seven participants stated that they used this compartment by habit.

Successful tasks. Four participants selected always (in all four tasks) the proper compartment. All of them selected Windows compartment for task 1. Note, these four participants have none of the measured demographic properties in common.

Results for task 1 (Search in local net). One participant decided to select the university compartment and twelve to select the Windows compartment.

Further findings. There was one participant who selected for task 2 the university compartment and did not leave this compartment while trying to conduct the remaining tasks. Correspondingly, she was not able to properly conduct task 3. She said that the university compartment is the only secure one and thus did not want to leave it again to avoid falling for an attack.

5.3 Group-Bottom

The Group-Bottom conducted the task on the test laptop displaying the TrustBar at the bottom of the screen. The different paths participants took throughout the study are also shown in Figure 6.

Results relevant for case - easy, privacy requiring situations (Task 2). The number of participants who were in the Windows compartment before task 2 (PIN entry w/o attack) is ten (see Figure 6). Nine of them switched to the university compartment to conduct this task (and did not try to access the university client application from Windows compartment). Five of them claimed that they did that for security reasons. The other four said that it is a habit for them to do university-related tasks in the university compartment. The one participant who tried to conduct this task in the Windows compartment received the error message. After he read the security indicators on the TrustBar, he switched to the university compartment and conducted this task there.

Results relevant for case - flexibility requiring situations (Task 3). The number of participants who were in the university compartment before task 3 (Search in Internet) is ten (see Figure 6). All of them switched to the Windows compartment to conduct this task (and did not try to access the Internet/search engine from the university compartment). One of ten participants stated that he switched for security reasons. The remaining nine participants stated that they used for such tasks the Windows compartment by habit.

Results relevant for case - difficult, privacy requiring (Task 4). The number of participants who were in the Windows compartment before task 4 (PIN entry w/ attack) is ten (see Figure 6). Six of them switched to the university compartment to conduct this task (and did not try to access the university client application from Windows compartment). Three of the six participants stated that they changed the compartment for security reasons and the other three claimed that it is a habit for them to do university related tasks in the university compartment. Four of the ten participants entered their PIN code in the fake university client application in the Windows compartment. All of them switched to the university compartment and conducted this task there after they received the error message. One of them said that he did not pay attention to in which compartment he was.

Successful task. Five participants selected always the proper compartment. These five participants have none of the measured demographic properties in common.

Results for task 1 (Search in local net). Three participants decided to select the university compartment and ten to select the Windows compartment.

Further findings. There were three participants who selected for task 1 the university compartment and did not leave this compartment while trying to conduct the remaining tasks. Correspondingly, they were not able to properly conduct task 3. One of them did not notice any security indicators on the TrustBar. He wanted to perform all tasks in the same compartment. The other two said that they saw the security indicators on the TrustBar, but they ignored them while trying to perform the tasks

5.4 Combining Numbers from Both Groups

Table 1 shows both the separated as well as the combined numbers for each of the tasks 2–4. The different paths that the participants took throughout the study are shown in Figure 6. For task 1, four participants decided to select the university compartment and 22 to select the Windows compartment. In total, nine of the 26 participants selected always the proper compartment.

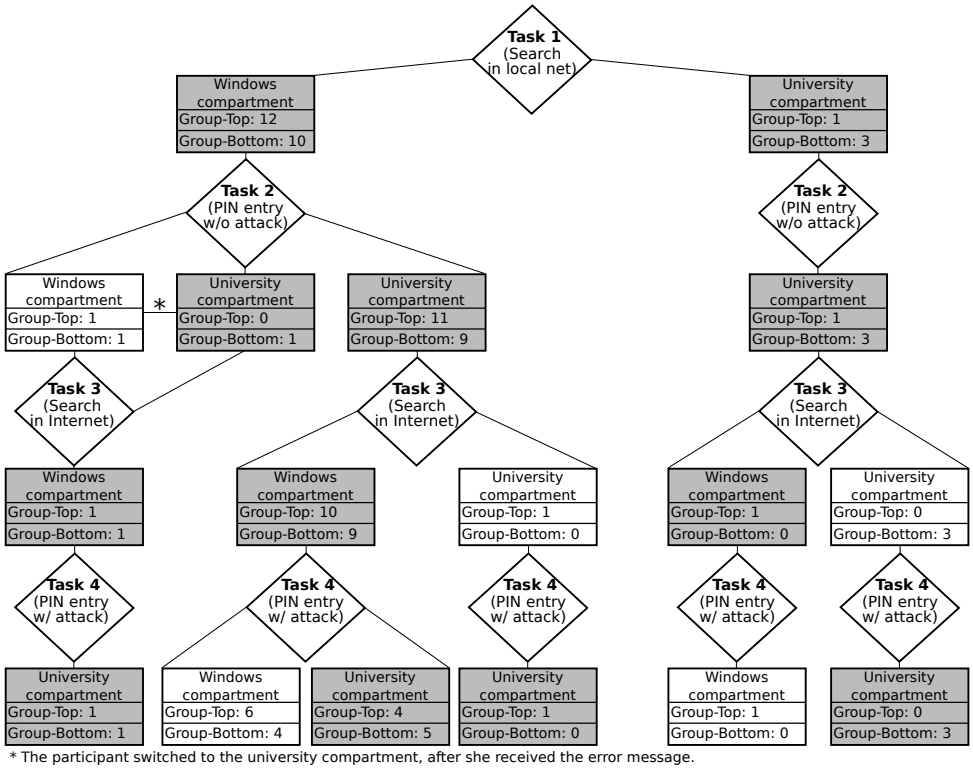

Figure 6: The number of participants who were in which compartment after each task.

	Task 2 (PIN entry w/o attack)	Task 3 (Search in Internet)	Task 4 (PIN entry w/ attack)
Group-Top	11 of 12 (91,7%)	11 of 12 (91,7%)	5 of 12 (41,7%)
Group-Bottom	9 of 10 (90%)	10 of 10 (100%)	6 of 10 (60%)
Total	20 of 22 (90,9%)	21 of 22 (95,4%)	11 of 22 (50%)

Table 1: Number of participants who decided to switch to the proper compartment

6. DISCUSSION AND CONCLUSION

Our results show that the participants have in general understood the meaning and the functionality of the different compartments. Group-Bottom performs slightly better—maybe because most of the participants are used to see status bars on the bottom of the screen with their own operating systems. While these results are very promising for the future deployment of Secure GUIs, the authentic-looking (but faked) university client application in the Windows compartment was convincing enough for in total eleven out of 22 participants. These numbers are better than the results for phishing web pages (90% of participants in [3], 92% of participants in [16]). Nevertheless, 50% falling for the attack of a fake university client application show that Secure GUIs or at least these two implementations do not yet protect users effectively and future research on appropriate user interfaces is necessary before deploying secure GUI techniques.

In future, new interfaces have to be researched. These include other positions of the TrustBar, maybe even individualized positions to enter/edit sensitive data (which is not known outside the corresponding compartment), new processes before entering sensitive data like a special keyboard combination to be pressed (similar to CRTL-ALT-DEL), and new hardware interfaces like additional LEDs. It looks currently very likely that different approaches for different applications are required, e.g., whether the compartment protects a PIN as in our scenario or sensitive (e.g. military) documents. Protecting the user more efficiently against authentic-looking but faked application will be a big challenge for the Secure GUI research area to provide adequate interfaces for desktop computers but it will be even more challenging for smartphones due to the smaller screen and the limited functionality.

While our participants performed very good for the first two tasks, there are some limitations with respect to the results. First of all, due to our project settings we only evaluated the interfaces with university students while studies with other groups of participants is left open. Furthermore, we only tested two compartments while usually users would likely need to deal with more compartments (also in our project they already have four on their own laptops). Correspondingly, studies with more compartments involved need to be conducted. Finally, we also tested only one type of application and one type of sensitive data—university client application and the students' PIN to access their data in the

university client application. In future, we plan to conduct studies containing other applications like online banking to study whether there is a difference in the results of complying for different applications and different application areas.

Regarding the tested tasks, we also noticed at the end of the study that it would have been technically possible to support participants in task 3 in case they try to access a search engine from the university department: Instead of showing a dialog just stating "access denied" one could recommend to go for another compartment. We will change the dialogs accordingly.

In total, one must say that there has been a lot of research on the technical aspects of Secure GUIs. However, to the best of our knowledge, this is the first time that researchers evaluate whether very common interfaces of Secure GUIs protect users effectively—which is not the case. It should also be noted that it is very difficult to run user studies in this area outside a project like ours because people need to have at least a basic understanding of the different compartments and have used such a system in order to then test whether they fail for fake applications or not.

7. REFERENCES

[1] W. A. Arbaugh, D. J. Farber, and J. M. Smith. A secure and reliable bootstrap architecture. In *Proceedings of the IEEE Symposium on Research in Security and Privacy*, pages 65–71, Oakland, CA, May 1997. IEEE Computer Society.

[2] A. De Luca, B. Frauendienst, M.-E. Maurer, J. Seifert, D. Hausen, N. Kammerer, and H. Hussmann. Does moodyboard make internet use more secure?: evaluating an ambient security visualization tool. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 887–890. ACM, 2011.

[3] R. Dhamija, J. D. Tygar, and M. Hearst. Why phishing works. In *CHI '06: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 581–590. ACM, 2006.

[4] J. Dumas and J. Redish. *A practical guide to usability testing*. Lives of Great Explorers Series. Intellect, 1999.

[5] J. Epstein. A prototype for Trusted X labeling policies. In *Proceedings of the Sixth Annual Computer Security Applications Conference (ACSAC)*, pages 221–230. IEEE, 1990.

[6] J. Epstein. Fifteen years after TX: A look back at high assurance multi-level secure windowing. In *ACSAC '06: Proceedings of the 22nd Annual Computer Security Applications Conference*, pages 301–320. IEEE Computer Society, 2006.

[7] J. Epstein, J. McHugh, H. Orman, R. Pascale, A. Marmor-Squires, B. Danner, C. R. Martin, M. Branstad, G. Benson, and D. Rothnie. A high assurance window system prototype. *Journal of Computer Security*, 2(2):159–190, 1993.

[8] G. Faden. Solaris Trusted Extensions: Architectural Overview. Sun Microsystems White Paper, Apr. 2006. `http://opensolaris.org/os/community/security/projects/tx/TrustedExtensionsArch.pdf`.

[9] N. Feske and C. Helmuth. A Nitpicker's guide to a minimal-complexity secure GUI. In *Proceedings of the 21st Annual Computer Security Applications Conference*, ACSAC '05, pages 85–94, Washington, DC, USA, 2005. IEEE Computer Society.

[10] T. Fischer, A.-R. Sadeghi, and M. Winandy. A pattern for secure graphical user interface systems. In *3rd International Workshop on Secure systems methodologies using patterns (SPattern'09), Proceedings of the 20th International Workshop on Database and Expert Systems Applications*, pages 186–190. IEEE Computer Society, 2009.

[11] T. Garfinkel, B. Pfaff, J. Chow, M. Rosenblum, and D. Boneh. Terra: a virtual machine-based platform for trusted computing. In *Proceedings of the 19th ACM Symposium on Operating Systems Principles (SOSP'03)*, pages 193–206. ACM, 2003.

[12] Green Hills Software Inc. INTEGRITY Real-Time Operating System. `http://www.ghs.com/products/rtos/integrity.html`, Nov. 2008.

[13] B. Lampson. Usable security: How to get it. *Communications of the ACM*, 52(11), 2009.

[14] C. E. Landwehr. Green Computing. *IEEE Security & Privacy*, 3(6):3, November/December 2005.

[15] B. Parno, J. M. McCune, and A. Perrig. Bootstrapping trust in commodity computers. In *Proceedings of the 2010 IEEE Symposium on Security and Privacy*, pages 414–429. IEEE Computer Society, 2010.

[16] S. E. Schechter, R. Dhamija, A. Ozment, and I. Fischer. The emperor's new security indicators. In *Proceedings of the 2007 IEEE Symposium on Security and Privacy*, pages 51–65. IEEE Computer Society, 2007.

[17] J. S. Shapiro, J. Vanderburgh, E. Northup, and D. Chizmadia. Design of the EROS trusted window system. In *Proceedings of the 13th USENIX Security Symposium*, pages 165–178. USENIX Association, 2004.

[18] R. Sherman, G. Dinolt, and F. Hubbard. Multilevel secure workstation. U.S. Patent 5,075,884, 1991. issued December 24.

[19] T. Whalen and K. M. Inkpen. Gathering evidence: use of visual security cues in web browsers. In *Proceedings of Graphics Interface 2005*, GI '05, pages 137–144, School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada, 2005. Canadian Human-Computer Communications Society.

[20] M. Wu, R. C. Miller, and S. L. Garfinkel. Do security toolbars actually prevent phishing attacks? In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, CHI '06, pages 601–610, New York, NY, USA, 2006. ACM.