User Perception of Risks Associated with IFTTT Applets: A Preliminary User Study

Bernardo Breve[1,*], Gaetano Cimino[1], Vincenzo Deufemia[1] and Annunziata Elefante[1]

[1]*Department of Computer Science, University of Salerno, Fisciano, Italy*

Abstract

Trigger-Action Platforms (TAPs) enable users to define rules that trigger device operations automatically. However, the execution of these rules can potentially create security risks for users. This paper presents a user study conducted to assess the validity of a classification model, which used Natural Language Processing (NLP) techniques to automatically classify Event-Condition-Action (ECA) rules according to security and privacy risks in TAPs, e.g., IFTTT. The study asked each user to evaluate 50 different IFTTT rules, named *applets*, classified as risky by the proposed model and provide answers to two specific questions designed to assess risk perception. The results confirmed that the proposed classification model offers an assessment of the risk associated with a rule in line with user opinion. Furthermore, highlighting the presence of security or privacy-related risk positively impacted users' willingness to avoid using risky applets.

Keywords

Usable Security and Privacy, Trigger-Action Platforms, Internet of Things, User Perception, Human issues and awareness,

1. Introduction

The vast spread of the Internet of Things (IoT) [1] has revolutionized the way we live our daily lives, turning home appliances, speakers, thermostats, and other devices into their smart variants, equipped with an Internet connection that enables them to collect and share information with other devices, leading the creation of factual ecosystems [2]. In an effort to enable all categories of users, including those with no technical knowledge, to take full advantage of the use of IoT devices, a number of platforms have emerged in recent years that make it easy for users to configure smart devices and define automation [3, 4]. The use of these platforms, named Trigger-Action Platforms (TAPs), allows the definition of interoperability behaviors between IoT devices through the creation of simple rules based on the Event-Condition Action (ECA) paradigm [5], i.e., specifying the event that triggers the automatism when a certain condition is met and the subsequent action that will be taken. Among TAPs, If-This-Then-That (IFTTT) affirmed itself in the last years as one of the most used ones, mainly thanks to the vast catalog

ITASEC 2023: The Italian Conference on CyberSecurity, May 03–05, 2023, Bari, Italy

[*] Corresponding author.

[†] These authors contributed equally.

bbreve@unisa.it (B. Breve); gcimino@unisa.it (G. Cimino); deufemia@unisa.it (V. Deufemia); anelefante@unisa.it (A. Elefante)

0000-0002-3898-7512 (B. Breve); 0000-0001-8061-7104 (G. Cimino); 0000-0002-6711-3590 (V. Deufemia); 0009-0001-7141-6105 (A. Elefante)

of rules (a.k.a. applets) that are available to the users and also for the considerable amount of services and devices that IFTTT supports.

Unfortunately, the IoT domain as well as technology in general can be characterized by several vulnerabilities which, if exploited by malicious individuals, can cause serious risks to the security of the IoT ecosystem and the privacy of users interacting with it [6, 7]. Sometimes, it may be the user who, inexperienced on security and privacy topics, introduces vulnerabilities himself through ECA rules [8, 9]. For instance, the rule in Figure 1 allows the user to automatically tweets anytime s/he enters a certain area. The user can decide to use this rule to notify his/her followers that s/he just arrived at the gym. However, this rule could pose an important security risk, as thieves can gain awareness concerning the users' absence by monitoring his/her routine through the tweets, helping them plan a break-in.

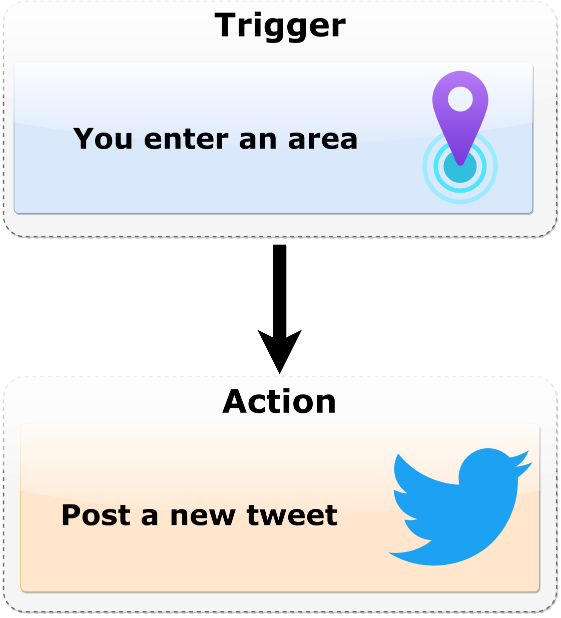


Figure 1: An example of ECA rule

In a previous publication, we addressed the problem of identifying security and privacy risks underlying the definition of ECA rules [10]. In particular, we proposed the application of Natural Language Processing (NLP) techniques to automatically classify ECA rules according to security and privacy risks. The application of NLP techniques allowed us to semantically analyze the triggers, actions, and natural language textual descriptions provided by the rule. The best model we considered, based on the Bidirectional Encoder Representations from Transformers (BERT) by Google achieved very high accuracy scores, with an average of 88%.

In addition to the empirical evaluations we have already conducted, we sought to ascertain whether the risks we identified could also be recognized and confirmed by the end-users who are the primary audience for our initial proposal. Thus, in this paper, we present a user study involving a group of 30 individuals. Specifically, we asked each user to evaluate 50 different IFTTT applets classified by our model proposed in [10]. For each applet, the user was asked to provide the answer to two specific questions designed to assess risk perception in view of the classification performed. The collected responses served a dual purpose of evaluating users' perception of the identified risk type as plausible and examining if reporting the risk could influence their decision to use or avoid the risky applets. The study findings were consistent with the empirical evaluations, indicating that our classification model accurately assesses applet-associated risks according to user perspectives. Moreover, bringing attention to potential security or privacy risks had a positive effect on users' willingness to avoid using unsafe applets.

The rest of the paper is organized as follows: Section 2 presents the main studies published in the literature assessing users' perceptions regarding security and privacy risks in IoT environments. Section 3 briefly summarizes the contribution related to the definition of the BERT-based classification model for identifying risks related to ECA rules. Section 4 presents the user study, discussing the evaluation setup and the obtained results. Finally, in Section 5, conclusions and planned future developments are drawn.

2. Related Work

The interaction with TAPs can sometimes pose serious risks both for the privacy of the user and/or the security of the smart environment [9]. In fact, rules created through such platforms define automatisms hiding unexpected behaviors that may go unnoticed by users. This becomes particularly evident if we consider the lack of technical knowledge the end-users have [7, 11].

The perception of users with respect to the privacy and security risks arising from rules defined through TAPs is a crucial aspect that needs to be considered. Various studies have been conducted to investigate this topic, aiming to understand how users perceive the potential risks of granting third-party access to their personal data and devices, as well as the measures they take to protect their privacy and security. Saeidi *et al.* conducted a study to investigate the implicit risks of using trigger-action platforms such as IFTTT in connecting smart-home devices and services [12]. They surveyed 386 participants on 49 smart-home IFTTT applets using a Mechanical Turk survey and found that users were generally not very concerned about using the rules, with the lowest level of concern being the most frequently selected answer. The study also identified the types of rules that elicited more concerns from users, which were those that involved acquiring, processing, or sharing location data. The authors suggest that nudging participants to think about different usage contexts led them to raise their concern scores. The study presented in [7] aimed to investigate the rationales behind smart home device purchases, homeowners' perceptions of privacy risks, and the measures taken to protect privacy from external entities such as device manufacturers, governments, Internet Service Providers, and advertisers. The study involved 11 semi-structured interviews with smart homeowners, and the analysis identified recurring themes. First, users' preferences for convenience and connectedness influenced their privacy-related behaviors in dealing with external entities. Second, users' opinions about external entities collecting smart home data were based on the perceived benefits of these entities. Third, users trusted IoT device manufacturers to protect their privacy but did not verify the implementation of such safeguards. Finally, users were unaware of the privacy risks from inference algorithms operating on data from non-audio/visual devices. The study's results suggest recommendations for device designers, researchers, and industry standards to match device privacy features with the expectations and preferences of smart homeowners. In [9], the authors analyzed 732 applets installed by 28 participants and their responses to survey questions to study the risks of real-world use of IFTTT. The study found that although public applets on IFTTT present a potential attack vector, most participants preferred creating their own applets. While participants did not express significant concerns about security and privacy risks from their use of IFTTT, they were aware of the possibility of such risks. Additionally, four participants reported experiencing applet-related harms or applets that did not function as expected. Overall, participants stressed the importance of security and privacy for their applets, and expressed concerns about applets triggering unintentionally, posting private information, spreading malware, or damaging smart-home devices.

The results suggest that TAPs should offer support to end-users in effectively managing and comprehending the security and privacy risks associated with creating trigger-action rules. To address this issue, specific efforts have been dedicated to developing ad-hoc solutions that enable end-users to identify and mitigate these risks [10, 13]. Moreover, it is crucial to ensure that users can fully comprehend the identified risks in order to make informed decisions. Consequently, various studies have proposed approaches for generating explanations that describe the causes of system instability [14, 15]. Such explanations aim to provide a clear and understandable account of the underlying technical concepts, as well as the potential consequences of certain trigger-action rules. These explanations may also help users to understand the trade-offs between privacy and functionality, and to make more informed decisions about whether or not to grant access to their personal data and devices. However, further research is needed to evaluate the effectiveness of these approaches, and to determine how they can be integrated into existing trigger-action programming tools.

3. Methodology

This section presents a comprehensive description of the sequential steps undertaken to produce a fully labeled dataset that encompasses ECA rules, followed by the process of training and evaluating a classification model that targets the identification of harmful rules. The proposed methodology was applied to a case study concerning the IFTTT platform.

3.1. System Overview

There are three fundamental phases involved in the process of constructing a classifier to identify harmful ECA rules:

- The first phase, named "*Data Labeling*", is designed to create labeled datasets for classification models. This is accomplished by defining the possible classes of risk for ECA rules, and their corresponding labels. Each ECA rule in the input dataset is then annotated with a suitable label using a semi-automatic labeling strategy that partitions the dataset into a small manually labeled subset and a larger subset that is automatically labeled using semi-supervised classification models.
- The second phase, named "*Model Training*", focuses on training the classification models using the labeled ECA rules dataset. NLP techniques are used to extract semantic infor-

mation from the textual components of the ECA rules, and a weighted loss function is applied to deal with the imbalanced nature of the training set.
- The last phase, named "*Model testing*", involves evaluating the performance of the classification models by inputting a set of manually labeled ECA rules and measuring their precision, recall, F1-score, and accuracy.

3.2. Data Labeling

This section introduces the dataset employed in training the classifier for identifying harmful applets and the process by which the applets were labeled. The applet labeling process comprised a dual approach of manual and automatic labeling, involving semi-supervised models and an ensemble strategy.

IFTTT Applet Dataset. The study was based on the dataset proposed by Mi *et al.* [16]. It was generated by researchers from Indiana University Bloomington, who conducted a web crawl of the IFTTT.com site over a period of six months, from November 2016 to May 2017. During this time, they collected a "snapshot" of the available applets each week, resulting in a dataset of over 300,000 unique applets, which totaled approximately 200 GB of data. The dataset contains essential information, including the applet name, description, trigger, trigger channel, action, action channel, and the number of users who have installed each applet.

The dataset underwent a data cleaning process to obtain a uniform dataset in language, and the LANGDETECT Python library was used to filter out applets not written in English. Applets without a name or description, or containing only numbers for these features, were discarded. After the data cleaning process, the resulting dataset contained 116,825 applets.

Categorization of IFTTT Applets According to Security and Privacy Risks. To categorize the potential damages that could be inflicted by an applet on the user, we referred to the work presented in [17], where potential damages were classified into four macro-categories. The first category, referred to as **Innocuous**, comprises applets that do not pose any harm or risks. The second category, labeled as **Personal**, includes applets that may result in the loss or compromise of sensitive data, which is solely due to the user's behavior. The third category, named **Physical**, involves applets that may cause physical harm or damage to goods, and the harm is external, i.e., inflicted by third parties. The fourth category, denoted as **Cybersecurity**, encompasses applets that may disrupt online services or distribute malware, and the harm is external as well.

We used the following classes for applet labeling, based on the considered macro-categories of risk: class *0* corresponds to Innocuous applets, class *1* to Personal damages, class *2* to Physical damages, and class *3* to Cybersecurity damages.

Manual Applet Labeling. The process of manually labeling the IFTTT dataset involved applying the majority method, whereby the first and second authors were responsible for the labeling process, and the third author mediated in cases of disagreement. This approach resulted in the labeling of 1,000 applets.

To increase the number of labeled applets, we developed a process for selecting additional applets to be manually labeled. This process involved creating a spreadsheet for each labeled applet, and sorting all unlabeled applets by their similarity to the labeled applet, using a combination of vector semantics and similarity functions. Specifically, we used SentenceBERT [18] to compute sentence embeddings for each applet and cosine similarity to compare the embeddings. We then manually reviewed each spreadsheet to identify and label applets that exhibited similar characteristics to those previously labeled, but differed in the triggers, actions, and/or channels involved. This process resulted in an augmented dataset consisting of 2,473 labeled applets.

Automatic Applet Labeling. After the manual labeling process, we developed a methodology that combines various semi-supervised learning models with an ensemble strategy. We utilized three distinct semi-supervised learning techniques, namely *Self Learning* [19], *Label Propagation* [20], and *Generative Adversarial Learning* [21]. *Self Learning* is a semi-supervised learning technique that uses a combination of labeled and unlabeled applets to train a model. This technique involves making predictions on unlabeled applets and treating those predictions as additional labeled applets to augment the existing labeled set. In contrast, *Label Propagation* propagates labels from a small set of labeled applets to a larger set of unlabeled applets. This is achieved by constructing a graph where the applets represent the nodes and the edges represent their similarity. Finally, *Generative Adversarial Learning* uses a generative model to generate synthetic applets similar to real ones and a discriminative model that learns to distinguish between them.

To obtain a single dataset from the three labeled applet datasets generated by the semi-supervised learning models, we employed an ensemble strategy. It consisted of a majority-vote scheme among the three semi-supervised models, in which applets were considered for inclusion in the final dataset with their respective class labels only if at least two models produced the same prediction. This approach provided us with more consistent labels for the evaluated applets.

3.3. Model Training

The dataset we constructed was characterized by an imbalance in the number of applets across classes, which poses a challenge in supervised classification [22]. In such scenarios, models trained on imbalanced data tend to classify input samples based on the majority class. To overcome this issue, we employed a weighted loss function that assigns different weights to each class based on the number of applets in the class. Notably, we assigned the minimum weight to class 0, which had the highest number of applets in the dataset.

We developed a BERT-based classifier to identify harmful applets by using applet information as textual features. The `BertForSequenceClassification` class of the TRANSFORMERS Python library was employed for training the classifier. This class corresponds to the BERT model with a single linear layer added for classification. We used the “bert-base-uncased” model, which is the base version of BERT with 12 transformer blocks, 768 hidden units, 12 self-attention heads, and lowercase letters.

3.4. Model testing

We conducted a series of experiments aimed at evaluating the effectiveness of the BERT-based model in classifying the different types of applet damage. To this end, a training dataset of 76,741 applets was assembled by employing the ensemble strategy combining the three sets of labeled applets generated with the semi-supervised learning models. The efficacy of the proposed approach and the quality of the labels produced by the ensemble strategy were validated by evaluating the model's performance on a test set comprising 2,473 manually labeled applets.

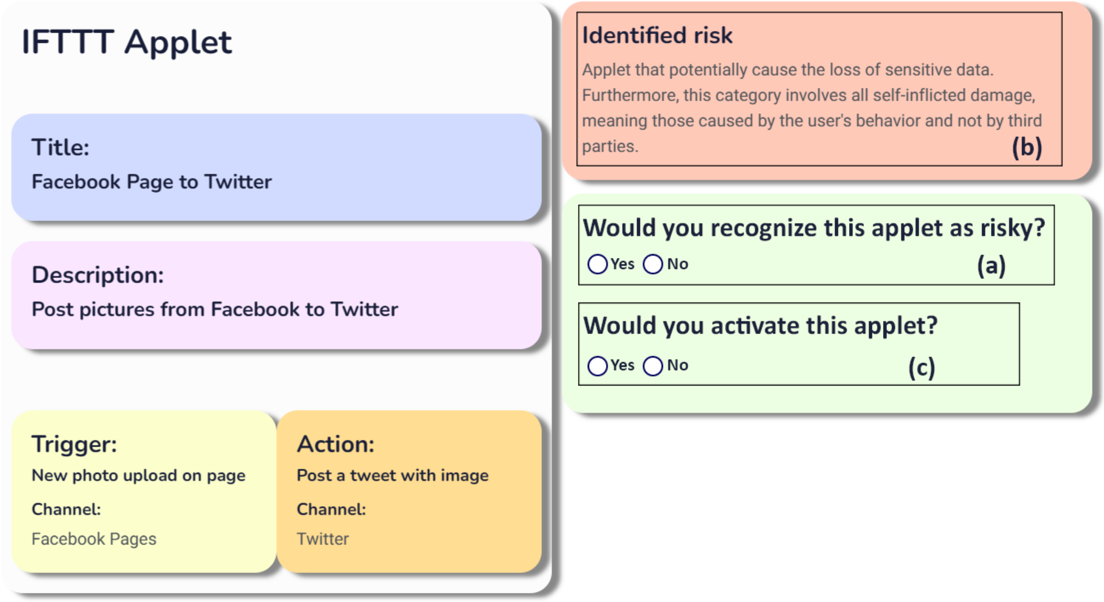
Table 1
Performances of the BERT model on the test set

Metric	class 0	class 1	class 2	class 3	WAvg
Precision (%)	89	87	88	87	88
Recall (%)	86	79	97	91	88
F1-score (%)	87	83	92	89	88
Accuracy (%)					88

Table 1 reports the values of accuracy, precision, recall, and F1-score achieved by the BERT model. Notably, the model attained a weighted average score of 88% across all metrics. Analysis of the results by class reveals that identifying class 1 applets is the most challenging task for the model. This difficulty arises from the slight differences in the context of rule execution that can cause errors by classifying class 1 applets as class 2 or 3. To illustrate, consider the applet "`Any new photo by me uploaded in a specific Google Drive folder, publish it on Twitter`". This applet falls under class 1 since it can lead to the unintentional sharing of sensitive or embarrassing photos. On the other hand, the applet "`Any new photo uploaded by anyone in a specific Google Drive folder, publish it on Twitter`" should be classified as class 3 due to the potential privacy risk of publishing a photo on the user's Twitter profile without their knowledge of who uploaded it. Similarly, the applet "`New tweet by me with a specific hashtag, turn off lights`" may be employed by a user to turn off lights with a goodnight tweet, but it can also trigger unintentionally in inappropriate situations, making it a class 1 applet. On the other hand, the applet "`New tweet by anyone in the area with a specific tag, turn on lights`" enables a user to determine if there are people who published a tweet in the zone, but its behavior may be jeopardized by third parties causing damage to the lights, making it a class 2 applet. These applets have subtle differences that make it challenging to classify class 1 applets compared to the other classes, which explains the lower performance of the model in this regard.

4. User study

This section describes the study we conducted to evaluate the user's perception of security and privacy issues related to IFTTT applets. In particular, we examined whether users are able to perceive the potential risks identified by the proposed BERT model as actual risks, and then evaluated whether such perception could potentially influence their decision to enable the



IFTTT Applet

Title:
Facebook Page to Twitter

Description:
Post pictures from Facebook to Twitter

Trigger:
New photo upload on page
Channel:
Facebook Pages

Action:
Post a tweet with image
Channel:
Twitter

Identified risk
Applet that potentially cause the loss of sensitive data. Furthermore, this category involves all self-inflicted damage, meaning those caused by the user's behavior and not by third parties.

Would you recognize this applet as risky?
○ Yes ○ No

Would you activate this applet?
○ Yes ○ No

Figure 2: An example of how an applet is presented to participants by our dedicated platform

corresponding applets. The research aimed to answer two questions, which were investigated through our experimental evaluation:

- *RQ1.* "Do users acknowledge the classification made by our model as an actual risk?"
- *RQ2.* "Can the identified risks impact on users' decision to activate the applet?"

To answer the two research questions, we recruited 30 volunteers (12 females) and designed a dedicated platform in charge of randomly selecting 50 IFTTT applets, contained within the dataset we released in [10]. Out of the applets selected for each test, 40 were belonging to one of three risk classes, *Personal*, *Physical*, or *Cybersecurity*, while the remaining 10 applets had been classified as *Harmless.* This distribution of applets had two advantages: firstly, it served as a control to prevent any pattern in user responses by randomly presenting 50 applets, some harmless and some harmful. Secondly, for harmless applets, open-ended responses were recorded on whether there was a discrepancy between the user's evaluation and the model's one, i.e., when the user interprets an applet as harmful but the model did not. The participants were selected from a mixed pool of users, consisting of individuals pursuing bachelor's and master's degrees in computer science as well as those from other academic disciplines, with a mean age of 28 years. Figure 2 shows an example of how applets are presented to the users and the related questions asked to them. Each applet was presented to the participants in terms of its title, description, trigger, and action components. Initially, the box describing the identified risk (b) for a certain applet is hidden from the user, as well as the second question (c), while the only operation a participant could perform is answering the first question (a). Once the users answer the first question, blocks (b) and (c) appear, and users are asked to confirm whether they would activate the applet in view of the identified risk. The evaluation process lasted on average 20 minutes per user.

To avoid answers that were biased over personal considerations outside the scope of our evaluation, we made it clear to the users to provide their answers only according to the risk and not to their personal needs. This is to avoid, for example, that the applet having the description "`Turn off my Philips Hue light anytime I leave the apartment`" would not be activated by a user, not because of the risk it implies, but because such a user did not have a Philips Hue light.

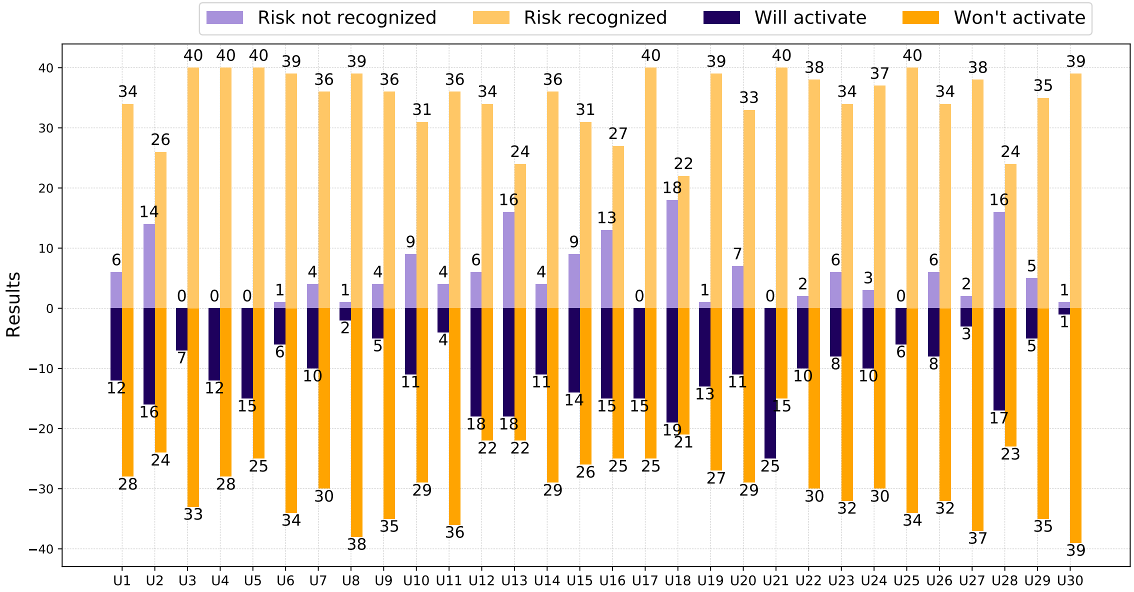


Figure 3: Summary bar plot of the results obtained from the user study

Figure 3 shows the results obtained from the user study, considering only the 40 risky applets presented to each user. Arranged on the x-axis are the 30 users who were involved in the experiment, while the y-axis collects the number of responses received for both the first question, the one about risk perception, and the second question, the one that asked about the user's confirmation to activate the applets.

In detail, the top part of the plot compares the responses obtained from the first question. While the left bar (the one in the light purple) shows the number of negative responses, i.e., the number of times in the user's opinion the applet presented did not pose a security or privacy risk. On the other hand, the bar on the right of each user shows the number of times a given user considered the type of risk identified by the model to be likely.

As for the lower part of the plot, the number of responses regarding the second question is reported. In this case, the left bar (the one in the dark purple) reports the number of times that, at the expense of the highlighted risk, the user would still decide to activate such an applet in their smart environment. Accordingly, the right bar highlights the number of times when the presence of an identified security or privacy risk leads the user to decide to avoid activating such an applet.

The results obtained are very promising and indicative. In fact, even at first glance, it can be seen that the number of times the risk identified by our model was deemed likely is very high, with an average of 34.7 out of 40. For many users (U3, U4, U5, U17, U21, U25), the totality of the applets they were asked to evaluate had security and privacy risks that conformed with what the model highlighted. In a few cases (U2, U13, U18, U28) it happened that the risk was highly not recognized by the participants, resulting in more than 10 negative responses. When questioned about the reason for such a high number of unrecognized risks, all of the above-mentioned users gave the same answer, namely that in their opinion the highlighted risk was related to scenarios so remote as to be almost unrealistic, reasoning that such applets would never come to pose any security or privacy risk.

Regarding the users' responses on whether they planned to activate the applets, this study revealed some variability in the responses, in fact it was observed that none of the users chose to refrain from activating all 40 applets that were presented for evaluation. It is worth noting, however, that even for this type of evaluation, users almost always opted not to activate the applets highlighted as risky. The only exception is user U21, for whom the number of applets he would not activate (15), turns out to be lower than the number of applets he would activate anyway (25). This result is even more surprising in light of the fact that U21 belongs to that group of users who instead indicated that all the applets that were submitted to him were likely to be risky. When questioned about this ambivalence in the answers, the participant responded as follows: "I agree that all applets are in some way risky, however, in my opinion, having them available would still be convenient. But now, knowing that they are risky, I would perhaps be more careful to avoid falling into problems (of security or privacy)".

On the other hand, with regard to users U2, U13, U18, and U28, who had been the ones who had most indicated the identified risk as "unlikely", the responses collected to this second question confirm their thinking. These users, in fact, are among those who would activate most of the applets presented to them. This result should not be surprising since in view of the fact that they do not consider many applets as risky, it is reasonable to assume that such applets can (and should) be activated either way.

As for the results concerning the 300 harmless applets we submitted to the users as a control, our results highlighted that 39 (13%) of them were instead judged as harmful by the participants. As mentioned above, for such applets we asked the participants to provide us with an open-ended answer, explaining why, in their own view, such an applet would be capable of causing a danger to the user. Table 2 shows some of the most relevant answers we received from participants and a short summary of the considered applet. In particular, some users e.g. U3, highlighted that the overuse of certain automation could lead to concerns since some services might not be capable of keep functioning as intended when involved in applets, such as the Evernote cloud space that could be filled up if provided with multiple notes. On the other hand, users such as U12 and U34 suggested that certain applets should be considered as harmful whether certain conditions apply, like when minors are involved or when there is a risk of fueling the spread of malicious software.

In conclusion, the experimental evaluations provided interesting insights concerning the 2 RQs posed upstream of this study. In fact, with respect to *RQ1*, the BERT-based harm classification model we presented in [10] provided risk classifications of IFTTT applets that were highly compatible with the thought of the users involved in the experiment, further demonstrating its reliability, already highlighted by empirical evaluations. Regarding *RQ2*, on the other hand, the results show that highlighting the presence of security and privacy risks related to IFTTT

Table 2
Opinions of users regarding certain applets categorized as harmful.

USER ID	APPLET SUMMARY	OPEN ANSWER
U3	Applet designed to create a new note in Evernote each time you add an item to the To Do List	May cause the cloud space provided by Evernote to fill up
U12	Applet with the title “Kids Education” designed to create a link post on Tumblr when a new post on Blogger	I think there are always serious risks when there are videos involving minors
U34	Applet designed to automatically create a link post on a Facebook Page when a top-rated app has gone free in the Apple App Store	Because if the top-rated app is a malicious app, you may be posting a dangerous app publicly on your profile, favoring its spread

applets positively influenced users not to activate the applets, or at least to weigh their use to avoid the occurrence of those risk scenarios. Finally, some harmless applets were instead identified as harmful by some participants, and some of their suggestions stressed the need to take a broader view with respect to how to assess risk.

5. Conclusions

In this study, we investigated the privacy and security concerns associated with applets activated through the IFTTT platform. We recruited 30 users to evaluate applets classified as risky by our previously proposed BERT-based classification model [10]. The results demonstrate the effectiveness of the model in ranking risky applets in accordance with user opinions. Furthermore, our analysis reveals that users can perform accurate risk assessments when provided with appropriate warnings. Specifically, we found that participants demonstrated a greater level of consideration towards specific applets, often opting to use the applet more consciously or not activate it to prevent potential risks. These findings suggest that risk assessment and management should be incorporated into the design and implementation of TAPs, with adequate warning mechanisms being provided to users. Future research could also investigate the effectiveness of additional strategies for enhancing user awareness and understanding of privacy and security risks associated with automated platforms.

6. Acknowledgments

This work has been supported by the Italian Ministry of University and Research (MUR) under grant PRIN 2017 “EMPATHY: Empowering People in deAling with internet of THings ecosYstems” (Progetti di Rilevante Interesse Nazionale – Bando 2017, Grant 2017MX9T7H).

References

[1] L. Atzori, A. Iera, G. Morabito, The internet of things: A survey, Computer networks 54 (2010) 2787–2805.

[2] B. L. R. Stojkoska, K. V. Trivodaliev, A review of internet of things for smart home: Challenges and solutions, Journal of cleaner production 140 (2017) 1454–1464.

[3] C. Ardito, P. Buono, G. Desolda, M. Matera, From smart objects to smart experiences: An end-user development approach, International Journal of Human-Computer Studies 114 (2018) 51–68.

[4] G. Ghiani, M. Manca, F. Paternò, C. Santoro, Personalization of context-dependent applications through trigger-action rules, ACM Transactions on Computer-Human Interaction (TOCHI) 24 (2017) 1–33.

[5] G. Desolda, C. Ardito, M. Matera, Empowering end users to customize their smart environments: model, composition paradigms, and domain-specific tools, ACM Transactions on Computer-Human Interaction (TOCHI) 24 (2017) 1–52.

[6] B. Breve, G. Desolda, V. Deufemia, F. Greco, M. Matera, An end-user development approach to secure smart environments, in: End-User Development: 8th International Symposium, IS-EUD 2021, Virtual Event, July 6–8, 2021, Proceedings, Springer, 2021, pp. 36–52.

[7] S. Zheng, N. Apthorpe, M. Chetty, N. Feamster, User perceptions of smart home IoT privacy, Proceedings of the ACM on human-computer interaction 2 (2018) 1–20.

[8] M. Surbatovich, J. Aljuraidan, L. Bauer, A. Das, L. Jia, Some recipes can do more than spoil your appetite: Analyzing the security and privacy risks of IFTTT recipes, in: Proceedings of the 26th International Conference on World Wide Web, 2017, pp. 1501–1510.

[9] C. Cobb, M. Surbatovich, A. Kawakami, M. Sharif, L. Bauer, A. Das, L. Jia, How risky are real users' IFTTT applets?, in: Proceedings of the Sixteenth USENIX Conference on Usable Privacy and Security, 2020, pp. 505–529.

[10] B. Breve, G. Cimino, V. Deufemia, Identifying security and privacy violation rules in trigger-action IoT platforms with NLP models, IEEE Internet of Things Journal (2022).

[11] Q. Wang, P. Datta, W. Yang, S. Liu, A. Bates, C. A. Gunter, Charting the attack surface of trigger-action IoT platforms, in: Proceedings of the 2019 ACM SIGSAC conference on computer and communications security, 2019, pp. 1439–1453.

[12] M. Saeidi, M. Calvert, A. W. Au, A. Sarma, R. B. Bobba, If this context then that concern: Exploring users' concerns with IFTTT applets, Proceedings on Privacy Enhancing Technologies 2022 (2021).

[13] F. Paci, D. Bianchin, E. Quintarelli, N. Zannone, IFTTT privacy checker, in: Emerging Technologies for Authorization and Authentication: Third International Workshop, ETAA 2020, Guildford, UK, September 18, 2020, Proceedings 3, Springer, 2020, pp. 90–107.

[14] B. Breve, G. Cimino, V. Deufemia, Towards explainable security for ECA rules, in: Proceedings of the 3rd International Workshop on Empowering People in Dealing with Internet of Things Ecosystems (EMPATHY '22), volume 3172 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2022, pp. 26–30.

[15] D. Xiao, Q. Wang, M. Cai, Z. Zhu, W. Zhao, A3ID: An automatic and interpretable implicit interference detection method for smart home via knowledge graph, IEEE IoT J 7 (2019) 2197–2211.

[16] X. Mi, F. Qian, Y. Zhang, X. Wang, An empirical characterization of IFTTT: ecosystem, usage, and performance, in: Proceedings of the 2017 Internet Measurement Conference, 2017, pp. 398–404.

[17] M. Surbatovich, J. Aljuraidan, L. Bauer, A. Das, L. Jia, Some recipes can do more than spoil your appetite: Analyzing the security and privacy risks of IFTTT recipes, in: Proceedings 26th International Conference on World Wide Web, ACM, 2017, p. 1501–1510.

[18] N. Reimers, I. Gurevych, Sentence-BERT: Sentence embeddings using Siamese BERT-networks, in: Proceedings International Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, ACL, 2019, pp. 3982–3992.

[19] D. Yarowsky, Unsupervised word sense disambiguation rivaling supervised methods, in: Proceedings 33rd Annual Meeting of the Association for Computational Linguistics, ACM, 1995, pp. 189–196.

[20] X. Zhu, Z. Ghahramani, Learning from labeled and unlabeled data with label propagation, Technical Report CMU-CALD-02-107, Carnegie Mellon University, 2002.

[21] D. Croce, G. Castellucci, R. Basili, GAN-BERT: Generative adversarial learning for robust text classification with a bunch of labeled examples, in: Proceedings of 58th Annual Meeting of the Association for Computation Linguistics, ACL, 2020, pp. 2114–2119.

[22] L. Wang, M. Han, X. Li, N. Zhang, H. Cheng, Review of classification methods on unbalanced data sets, IEEE Access 9 (2021) 64606–64628.