Biometric Identification System based on Object Interactions in Internet of Things Environments

Klaudia Krawiecka, Simon Birnbach, Simon Eberz and Ivan Martinovic
Department of Computer Science, University of Oxford
United Kingdom
Email: firstname.lastname@cs.ox.ac.uk

***Abstract*—Attributing interactions with Internet of Things (IoT) devices to specific users in smart environments is extremely important as it enables personalized configurations and access control. This requirement is particularly stringent when it comes to parental control measures designed to protect children from contact with dangerous machinery or viewing materials that are inappropriate for their age. To this end, we show that naturally occurring interactions with objects in smart environments can be used as a behavioral biometric in order to identify users. The heterogeneous nature of smart devices enables the collection of a wide variety of inputs from such interactions. In addition, this system model allows for seamless identification, without the need for active user participation or rearrangement of the IoT devices.**

We conduct a remote study taking place in six households composed of 25 participants. We demonstrate that our system can identify users in multi-user environments with an average accuracy of at least 91% for a single object interaction without requiring any sensors on the object itself. This accuracy rises to 100% when six or more consecutive interactions are considered.

I. Introduction

Everyday objects and appliances are being enriched with heterogeneous sensors to facilitate daily activities and provide a better user experience. Such devices are fundamental building blocks of modern Internet of Things (IoT) environments. By the end of 2022, smart homes are projected to surpass the 84 million mark in Europe alone [15]. As the number of such environments grows, so does the number of devices owned by households. The variety of data collected by such devices not only contributes to providing more customized services, but can also be used to protect such environments.

Despite the growing number of smart environments, they often do not have any built-in security controls. Due to the limited interfaces of IoT devices, most of the known safeguards cannot be easily implemented. However, the attribution of certain activities performed with these devices is necessary to limit unauthorized use. For example, parents may want to control their children's use of devices such smart thermostats, TVs, or other appliances [20]. Therefore, a smart home should be able to distinguish who is interacting with these objects at a given moment to decide whether certain actions can be safely executed. While many existing solutions focus on the recognition of specific activities and gestures performed by the users of IoT devices [14, 2, 16, 12, 11], in this case, the attribution of those activities to specific household members is desired. Some studies look at identifying users using various on-device sensors [5, 19]. However, these approaches assume the presence of cameras or other specific types of devices in the environment. This is not only impractical but also does not use the full potential of heterogeneous environments. Moreover, most such systems actively involve users in the identification process, which negatively impacts their usability.

In this paper, we propose a system that seamlessly identifies users based on their interactions with smart objects. Our system models physical interactions using the most common sensors in smart environments, such as inertial measurement units (IMUs) and microphones [13]. It operates in three configurations: On-object, Off-object, and Combined. An On-object configuration uses on-device sensors to extract features, whereas in Off-object, only sensors from nearby objects are used. In the third configuration, the system relies on both on-device and co-located sensors. We show that some types of interactions are more distinctive than others and co-located objects contribute to higher accuracy of identification.

To validate our approach, we collected samples of multiple user interactions from six real households with 25 participants in total. To achieve this during the Covid-19 pandemic, we designed a set of experiments that were carried out remotely, without the need for researchers to be present in the participants' environments. Overall, based on our analysis, the system achieves high identification accuracy, regardless of the size of the household or the environment configuration.

We make the following contributions in the paper:

- We build a prototype of a novel identification system based on physical interactions with smart objects.
- We create a dataset consisting of sensor data from 38 devices across six households involving 25 participants.
- We make all data and code needed to reproduce our results available online. [1]

II. Background & Related Work

In this section, we cover the most common approaches and methods that translate users' physical interactions with surrounding objects into biometric features. Multi-sensor fusion techniques take advantage of the heterogeneous nature of sensors embedded in such objects to improve activity recognition. Our system is based on the assumption that attributing various activities to specific users will enable their seamless

[1] https://github.com/ssloxford/BIS-IoT

identification, which will not only significantly improve user experience but also security in smart environments.

A. Multi-sensor fusion

Multi-sensor fusion systems focus on identifying specific actions using diverse types of inputs supplied by a variety of sensors. Instead of relying on one modality, signals are extracted from different sources [1]. Most of the research in this area has been developed to aid Human Activity Recognition (HAR). Some HAR systems recognize certain types of body movements by analyzing objects' orientation and acceleration [14, 2, 16]. Others precisely identify tasks such as cooking or cleaning the house by applying sensor fusion [12, 11]. Using multiple sensor modalities improves the activity recognition performance and creates a better reference frame [9]. A variety of techniques, including different ensemble methods, can be used to combine heterogeneous modalities [1].

B. Interaction-based biometric systems

The type of biometrics based on the physical interactions of users with objects around them, such as smartphones, smart watches or other mobile platforms, attempts to fully exploit the heterogeneous nature of built-in sensors. Such systems profile users by collecting relevant information about their unique movements and gestures when they perform certain tasks, including typing, clapping, walking, and so on [8, 4, 3, 17].

A similar approach can be applied to IoT environments. However, instead of a single device, the entire infrastructure can be used to capture these interactions. In this domain, SenseTribute [10] is the closest to our work. It is an interaction-based occupant identification system that extracts signals from accelerometers and gyroscopes. It ensembles and clusters user activities with supervised and unsupervised machine learning techniques. Our work extends SenseTribute in several ways by considering more sensor modalities, several households instead of just one, and an increased number of users. More importantly, our proposed system is able to employ co-located sensors, which are not located directly on objects, to fingerprint users. This enables not only significantly improved identification performance in comparison to SenseTribute (c.f., Section VI), but also makes it easier to deploy in existing smart homes, as it uses the natural locations of smart devices without restrictions on their placement. It also allows us to identify interactions with objects that do not have any sensors of their own.

III. System Design

The main goal of this work is to introduce an intuitive and seamless identification method for smart environments. User identification in smart environments is necessary to enable user-specific configurations, track access or usage patterns (to improve user experience), or implement parental controls. Smart environments should be able to fingerprint users during their usual interactions with smart devices. The wide range of sensor types currently used in smart homes enables systems to profile household users and distinguish them with high accuracy. This is made possible through the input from sensors of nearby devices already present in the smart home as the user interacts with the environment during their daily activities.

In short, the system should fulfill the following design goals:

1) The system should provide a *seamless* user experience and should not require the user to change their normal behavior, such as having to perform explicit actions or gestures.
2) The system should have *low friction* for the user. This means that the system must have a high identification accuracy to avoid the need for user intervention due to misidentification by the system.
3) The system should not have specific configuration requirements with respect to sensor placement. Instead, it should be applicable to existing deployments. In particular, the system should not require that every object used for identification be equipped with sensors, but rather allow the use of *co-located devices*.

A. System overview

Our system is applicable to smart environments where everyday objects, such as refrigerators or drawers, have been augmented with smart devices. Users of such environments can monitor the states of these smart objects, have access to enhanced functionality, and interact naturally with them during their daily activities.

Typical activities in smart environments involve several interactions with smart objects. For example, the user may first open the dishwasher to start unloading it. Then they open various cabinets and drawers to put away clean dishes. These interactions have physical effects on the environment that can be picked up by nearby sensors, and—as each type of interaction is performed in a unique manner by different people—they can provide a behavioral fingerprint that allows our system to distinguish and identify different members of a household. Moreover, sequences of these interactions can be used to further increase system performance. We focus on a single user being identified at a time, and do not consider multiple concurrent users as a simplification.

Our system does not require the object that is part of an interaction to be equipped with sensors itself. In fact, it is often beneficial for the system to consider sensors on other close-by objects. In our work, we therefore consider three possible deployment configurations: using only sensors that are fitted to the object in question (ON-OBJECT), only data from co-located sensors (OFF-OBJECT), or a combination of both (COMBINED).

B. Adversary model

As our proposed system identifies the person that interacted with physical objects, we consider attackers to have physical access to the house and the devices therein. Due to this assumption, attackers are considered to be relatively benign (e.g., a child wanting to access restricted TV channels or order their favorite food using the smart fridge). As a result, we exclude more sophisticated attacks such as modifying sensor firmware, tampering with network communication, spoofing sensor data or poisoning the training phase. Instead, we assume the attacker will interact with devices in the normal manner and be successful if they are incorrectly identified as an authorized user (i.e., a zero-effort attacker). Note that our accuracy metric is actually stricter than this, as a child being misidentified as another unauthorized child is considered a wrong classification but would not lead to incorrectly denying or granting access.

Another type of attack that could be relevant but is not evaluated in this paper is the imitation attack. Since we assume that an adversary has physical access to the IoT infrastructure, they can attempt to mimic the behavior of the legitimate user. For instance, children may try to imitate the gestures of their parents while they interact with the smart devices.

IV. Experimental Design

During our experiment, we gather data in real-world homes of various household sizes to study the uniqueness of physical interactions in multi-user environments. In order to emphasize the real-world nature of the experiment, we do not restrict participants in the number and type of appliances they use during the study. This ensures that our setup reflects the makeup of real households, rather than an artificial lab setting. This research project has been reviewed by the responsible research ethics committee at our university and has received formal approval (ref. CS_C1A_20_014-1).

A. Data collection

To study the use of smart home interactions for user identification, we gathered data using types of sensors that can typically be found in modern smart homes. The sensors used in our experiment are magnetic contact switches (which can detect open/close events similar to door/window contact sensors), USB microphones (which are only used to measure sound pressure levels), as well as ICM20948 IMUs (which consist of an accelerometer, a gyroscope, and a magnetometer). These sensors are mounted on and controlled through ten Raspberry Pis. We use these custom Raspberry Pi sensor boards because raw sensor data in commercial smart devices is typically inaccessible for developers. The sensor boards are deployed in places where equivalent smart devices are already common in most smart environments, such as home appliances or kitchen furniture. The ground truth for events related to interactions with smart objects—such as the opening of a kitchen drawer or the closing of a fridge door—is provided by the aforementioned magnetic contact switches. Figure 1 shows one of the sensor boards deployed in the kitchen of a study participant.

The Raspberry Pis securely stream the data to a remote server through a wireless hotspot provided by a smartphone. Additionally, the devices store a local backup of the measurement data. The smartphone further gives the user access to an app that guides the user through the experiment, labels and timestamps each run of the experiment, and synchronizes the time of all deployed Raspberry Pis.

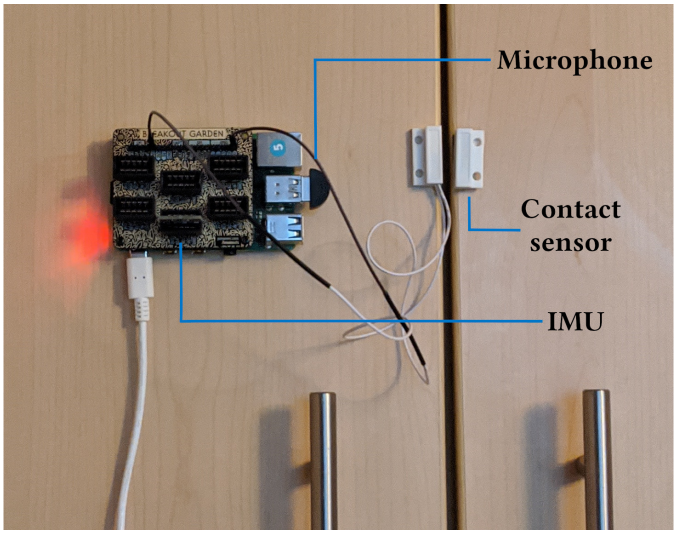

Fig. 1: A Raspberry Pi mounted on a kitchen cupboard door. It is outfitted with a magnetic contact sensor (to detect when the cupboard is opened or closed), a USB microphone (to measure sound pressure levels), and an IMU (to record acceleration, gyroscopic motion and orientation data).

B. Recruitment of participants

Due to the ongoing Covid-19 pandemic, we could only collect data from people from within their own households. Therefore, we advertised the experiment in our department as well as among friends and family in order to find households where several members of the household were willing to participate. We recruited a total number of 25 people including two children across six households. A detailed overview of the composition of the different households and the devices available in each can be found in Table I.

TABLE I: Household overview. The number of people for H6 includes two children.

Household	H1	H2	H3	H4	H5	H6
Number of people	3	2	6	4	3	7
Number of objects	8	10	4	5	4	7

TABLE II: Device location and interaction type breakdown. Numbers in parentheses show the number of related devices.

Device location	H1	H2	H3	H4	H5	H6
Fridge/Freezer	✓(2)	✓(2)	✓(2)			
Cupboard	✓(2)	✓(3)	✓(2)	✓(2)	✓(2)	✓(3)
Floor	✓(2)					
Pull-out drawer		✓(1)		✓(1)		✓(1)
Microwave door	✓(1)	✓(1)		✓(1)	✓(1)	
Oven door		✓(1)		✓(1)		✓(1)
Dishwasher	✓(1)	✓(1)			✓(1)	✓(1)
Washing machine		✓(1)				
Bread box						✓(1)

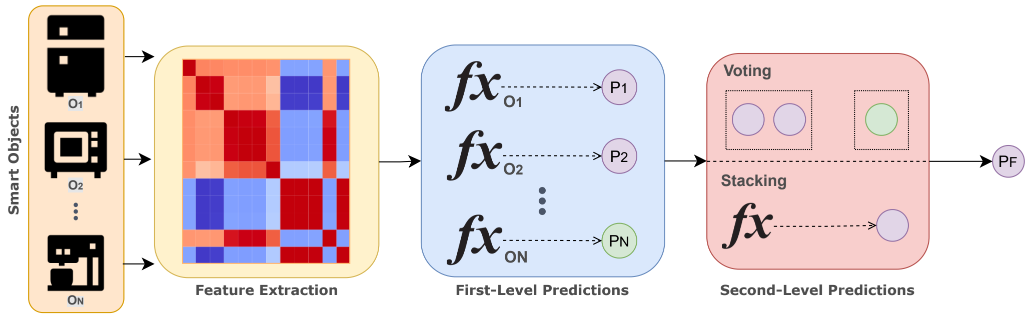

Fig. 2: The diagram provides an overview of the processing pipeline of an ensemble learning system. This system extracts relevant features from interactions with smart objects O_1 to O_n and supplies them to their base-classifiers. Then, the first-level predictions P_1 to P_n are fed into a meta-classifier (i.e., a voting or stacking classifier) that computes the final prediction P_F.

C. Remote household environments

Following Covid-19 regulations, our experiment is conducted remotely in the kitchens of the participants. The participants are given a set of our Raspberry Pi sensor boards and they have to set up the devices themselves, according to the provided step-by-step user manual. The number of used devices varies between households depending on the size of the kitchen, ranging from six to ten devices. Each of these devices corresponds to one object interaction (e.g., opening and closing the fridge door). The participants in each household choose these interactions themselves based on their kitchen layout. A breakdown of where the devices were installed for each household is given in Table II. For each of these interactions, one Raspberry Pi device is fitted to the corresponding monitored point of interaction. The only exceptions to this are the devices mounted on the floor in household H1. These devices do not provide any ground truth for device interactions but are instead used to give auxiliary sensor data to other interactions based on the gait characteristics of the user as they move around the room.

During every run of the experiment, participants perform each interaction exactly once. Each participant completes 20 runs of the experiment, resulting in a total of 3000 runs for our 25 participants across the six households.

V. Methods

A physical interaction I initiated by the user U with an object O is modeled as a time-series $I = \{X_1, X_2, \ldots, X_z\}$, where X_t is a signal (i.e. a vector of sensor values) captured by on-device sensors at time t. Such signals originate from various heterogeneous sensors, including microphones, accelerometers, gyroscopes, and magnetometers. To illustrate this, we will take an example of a smart microwave equipped with Inertial Measurement Units (IMUs). This smart object can collect acceleration values as vectors of $\langle a_x, a_y, a_z \rangle$ when the user opens or closes its door. Because of the heterogeneous nature of the smart objects, we expect that the length of such vectors as well as their types will be diverse.

Figure 2 shows the overview of the processing pipeline of our identification system. When users interact with smart objects $O_1 - O_N$, on-device and co-located sensors extract features and supply them to weak-learners that compute first-level predictions $P_1 - P_N$. These predictions are then aggregated by the meta-classifier that decides on the final prediction P_F in the second-level prediction layer.

A. Preprocessing

Each interaction I captured by a smart object is timestamped by its contact sensor. We denote the start and end of I as t_0 and t_1. Our system segments the signals by the values of $t_0 - 1$ and $t_1 + 1$ before extracting features in the next phase. I is composed of values characteristic of z sensor types. As a result, for each I, a set of corresponding matrices $M = S_1, S_2, ..., S_z$ exists that contains vectors of different sensor values X_i between $t_0 - 1$ and $t_1 + 1$. The number of columns for a single matrix is determined by the sensor components (e.g. acceleration axis). For a smart object with three on-device sensors and two components per each sensor, three such matrices with six columns will be generated. They are then passed as input to the feature extraction function.

B. Feature extraction & selection

Depending on the system configuration, M may be retrieved from the smart object the user interacted with (On-object), the objects in proximity (Off-object), or a combination of both (Combined). The features are computed from each column of S_i, which contains sensor values extracted between $t_0 - 1$ and $t_1 + 1$ to account for signals that originate from the starting and ending movements of a physical interaction. The system extracts features from the time and frequency domains, including: mean, median, standard deviation, variance, kurtosis, skewness, shape factor, absolute energy, mean of central approx. of 2^{nd} derivative, mean and sum of absolute change, peaks, and Fourier entropy. To protect the privacy of the experiment participants, these features are extracted from sound pressure levels (SPLs) instead of actual audio recordings. Next, the system performs mutual information-based feature selection [6]. For each O_i, our system ranks the features and then selects the top 20.

C. Training

During the system's training phase, each user labels their interactions with their identity after they have authenticated themselves (e.g., via a setup PIN) to avoid mislabeled data and training data poisoning. These interactions during the training phase have to be performed individually (one user at a time). In our experiment, this labeling was implemented through a smartphone app (see Section IV-A). Each feature vector (see Section V-B) generated for an object interaction during training is then labeled with the user identity provided. To evaluate our system, we split the total dataset into training and test data using 10-fold cross-validation. This ensures each sample (object interaction) is used for testing exactly once. The training data generated by this process is then used to train the classifiers as described in the following subsections.

Although we gave participants the freedom to experiment over a few days, the vast majority did so on the same day, which may not reflect possible changes in their behavior over time. In practice, these changes can be accounted for via periodic retraining and template updates. We leave exploration of time stability for future work.

D. Ensemble learning

The main reason for using ensemble techniques in our system is the varying effectiveness of the classification of the interactions with individual smart objects and their sensors. Therefore, to boost the performance of various first-level classifiers, we use two ensemble algorithms. These algorithms create second-level predictions based on the results obtained in the previous step.

1) Voting: Our system implements hard voting [7] as a baseline ensemble method. This technique aggregates the predictions of weak-learners and then selects the class that received the majority of votes. In our case, each smart object implements a weak-learner that predicts a class given a set of features extracted by its embedded sensors. These predictions are fed into a meta-classifier that, based on the most frequently recurring class, outputs the final prediction $P_F = \mathrm{mode}(P_1, P_2, \dots P_n)$.

2) Stacking: In comparison to voting, stacking uses another machine learning algorithm as a meta-learner. This classifier is trained using predictions of weak-learners as its features [18]. Stacking can improve the system performance because it learns which smart objects identify users better and discards the objects that are less accurate. Our system uses the Random Forest algorithm as a stacking meta-classifier.

VI. Evaluation

In this section, we evaluate the identification performance of our system and compare it to SenseTribute [10]. For each household, we obtain a different dataset. The number of samples in each dataset varies based on the number of household members that participated in our experiment. The decomposition of smart objects and members in each household is described in Section IV.

TABLE III: Comparison of average accuracy results from individual objects between different households given On-object, Off-object, and Combined configurations. The numbers in parentheses indicate the number of household members in each household.

	H1	H2	H3	H4	H5	H6
No. members	(3)	(2)	(6)	(4)	(3)	(7)
On-object	0.97	0.99	0.93	0.92	0.97	0.88
Off-object	0.96	1.00	0.96	0.89	0.98	0.91
Combined	0.95	1.00	0.99	0.90	0.98	0.93

Table III presents the averaged accuracy scores in different households over individual objects. The results are shown for On-object, Off-object, and Combined configurations. Apart from two households, the Combined configuration exhibits the best performance and On-object the lowest. The On-object configuration is particularly low for $H6$, which can be explained by the placement of some sensors on the semi-automated cupboard, which resulted in the loss of some vital interaction data for this object. In comparison to SenseTribute, which achieves only 74% average on-object accuracy in a 5-person environment, our system achieves a higher score of at least 88% in more populated environments, resulting in a 14% improvement in identification performance over the SenseTribute system. Interestingly, relying solely on co-located sensors (Off-object) to capture the unique characteristics of the interaction is sufficient to achieve high identification results. Overall, we observe that including the data from both on-device and co-located sensors benefits the system the most. However, we also note that this depends on the physical arrangement of the objects in the house as well as their spacing. Unlike in other households, there were greater distances between objects in $H1$ and $H4$. This confirms our hypothesis that the physical setup will determine which system configuration is better for a given smart environment. The main benefit of our system is that it is able to leverage sensors of co-located devices. If co-located sensors are taken into account, the minimal identification accuracy in $H6$ can be improved by 3% and 5% for Off-object and Combined settings, respectively. Aside from significantly improving the identification results compared to the scenario used by SenseTribute—where only sensors on the device that is being interacted with are considered—this also makes our system more easily deployable, as we do not require that the sensors used for identification are located on the device in question.

Table IV shows the accuracy scores for nine types of common household objects in the Combined configuration. This provides a more detailed view of the performance based on various kinds of interactions with objects across all households. Since the experiments took place in the participants' houses and each had a different set of objects, we could not always ensure that the same types would be used. Therefore, the classification performance for each object type is calculated with different sample sizes, depending on the household. The interactions with microwave doors appear to have the

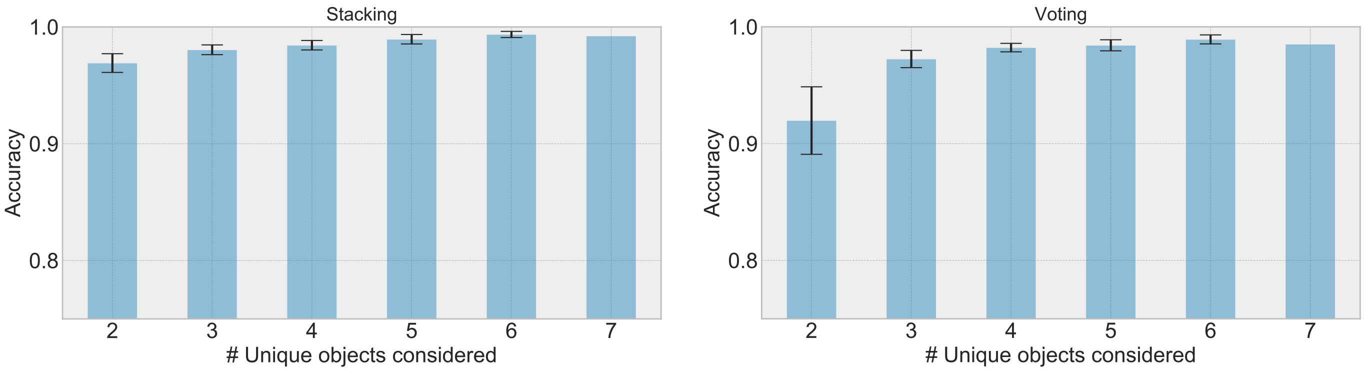

Fig. 3: Averaged accuracy scores of different ensembles of unique objects in $H6$ for two meta-classifiers and the COMBINED configuration. Each such ensemble is trained and tested separately, then the scores are averaged across the ensembles of the same type (e.g., pairs, triples of unique objects). The 95% confidence intervals are given for each ensemble. The y-axis cutoff range is between 0.75 and 1.00.

TABLE IV: Device interaction type and accuracy breakdown for individual devices using the COMBINED configuration.

Device location	H1	H2	H3	H4	H5	H6
Cupboard	0.96	1.00	0.99	0.95	0.98	0.91
Microwave door	0.89	1.00	-	0.80	0.98	-
Dishwasher	0.96	1.00	-	-	0.98	0.95
Fridge/Freezer	0.97	1.00	0.98	-	-	-
Pull-out drawer	-	1.00	-	0.90	-	0.96
Oven door	-	1.00		0.96	-	0.94
Floor	0.96	-	-	-	-	-
Washing machine	-	1.00	-	-	-	-
Bread box	-	-	-	-	-	0.92

TABLE V: Comparison of accuracy results between different households. Second-level predictions are generated using all the available objects in the household in the ON-OBJECT configuration and the Random Forest meta classifier. The numbers in parentheses indicate the number of household members in each household.

	H1	H2	H3	H4	H5	H6
No. members	(3)	(2)	(6)	(4)	(3)	(7)
Voting (V_{OT})	1.00	1.00	0.96	0.99	1.00	0.98
Stacking (S_{TC})	1.00	1.00	0.98	1.00	1.00	1.00

lowest classification performance, but still, achieve an average accuracy of 80% in a 4-person household. Typically, the COMBINED configuration improves the identification performance of objects with a low accuracy of interaction classification; however, as we mentioned above, in $H1$ and $H4$, the objects were far apart. Regardless of the size of the household, the average accuracy scores for interactions with other objects exceed 90%. In general, we observe that the interactions with the objects considered in our experiment are quite distinctive, which explains the good performance of the meta-classifiers.

The comparison between the identification performance of voting (V_{OT}) and stacking (S_{TC}) meta-classifiers across different households is summarized in Table V. The accuracy scores are calculated for ensembles of all objects in the ON-OBJECT configuration. We decided to present the worst-performing configuration to highlight the benefit of including more interactions. The table reveals that S_{TC} performs better. In this setting, the system learns which objects have better classification performance and relies on their predictions, rather than considering all. Unlike stacking, voting assigns equal weights to all base-classifiers; thus, they equally contribute to the final prediction. This could be further improved by assigning smaller weights to objects that exhibit worse performance. As expected, the performance of the two meta-classifiers begins to diverge as the number of household members increases. For instance, in households with 2–4 members, there is almost no difference in terms of performance of two meta-classifiers. However, for 6 and 7 people, the difference between the performance of V_{OT} and S_{TC} is 2%. Overall, the stacking meta-classifier outperforms the voting meta-classifier in terms of the identification performance. While this may also depend on the physical and behavioral differences between particular users (e.g., family members versus a group of students), we expect that in other households this would yield similar results. In comparison, SenseTribute reaches a maximal identification accuracy of 96% in their single five-person household setup when multiple objects are considered.

Figure 3 shows the average performance of different ensembles of unique objects (inc. 95% confidence intervals) in $H6$ in the COMBINED configuration. We focus on this household as it was the largest one in our experiment. The identification performance of the system, regardless of the selected meta-classifier, gradually improves when it considers more objects. Overall, S_{TC} achieves higher accuracy scores than V_{OT}. This difference is especially evident when considering two object interactions. The system that uses a stacking meta-classifier offers 5% improvement in the identification performance. The reason is, as explained earlier, that V_{OT} assigns equal weights to all base-learners. However, the gap between the ensembles of different size slowly decreases as more objects are included. Thus, the system can identify the users with high accuracy considering fewer object interactions.

VII. Future Work & Conclusions

In this paper, we propose an identification system based on physical interactions with smart objects. Rather than relying solely on built-in sensors, this system uses co-located sensors, placed arbitrarily in a smart environment.

We conducted an experiment with 25 participants in 6 different households, which showed that regardless of the size of the household or a specific household setup, users can be distinguished with high accuracy. Our proposed system is able to identify users with an accuracy of at least 91% on average if just one interaction is being recorded by the co-located sensors. This identification accuracy increases to 100% when multiple interactions are considered by the system.

These encouraging results indicate the potential use of this behavioral biometric for authentication purposes, which we intend to investigate in our future work. We publicly release our dataset and the code needed to reproduce our results.

Acknowledgments

The authors would like to thank Mastercard as well as the PETRAS National Centre of Excellence for IoT Systems Cybersecurity, which has been funded by the UK EPSRC under grant number EP/S035362/1 for financial support for this work. They would also like to thank the anonymous reviewers for their valuable comments.

References

[1] Antonio A Aguileta, Ramon F Brena, Oscar Mayora, Erik Molino-Minero-Re, and Luis A Trejo. Multi-sensor fusion for activity recognition—a survey. *Sensors*, 19(17):3808, 2019.

[2] Ashraful Alam, Anik Das, Md Tasjid, Ahmed Al Marouf, et al. Leveraging sensor fusion and sensor-body position for activity recognition for wearable mobile technologies. *International Journal of Interactive Mobile Technologies*, 15(17), 2021.

[3] Shurook S Almohamade, John A Clark, and James Law. Behaviour-based biometrics for continuous user authentication to industrial collaborative robots. In *International Conference on Information Technology and Communications Security*, pages 185–197. Springer, 2020.

[4] Buriro Attaullah, Bruno Crispo, Filippo Del Frari, and Konrad Wrona. Hold & sign: A novel behavioral biometrics for smartphone user authentication. 05 2016.

[5] Alex Barros, Denis Rosário, Paulo Resque, and Eduardo Cerqueira. Heart of iot: Ecg as biometric sign for authentication and identification. In *2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*, pages 307–312. IEEE, 2019.

[6] Mario Beraha, Alberto Maria Metelli, Matteo Papini, Andrea Tirinzoni, and Marcello Restelli. Feature selection via mutual information: New theoretical insights, 2019.

[7] Saso Džeroski and Bernard Ženko. Is combining classifiers with stacking better than selecting the best one? *Machine learning*, 54(3):255–273, 2004.

[8] Elakkiya Ellavarason, Richard Guest, Farzin Deravi, Raul Sanchez-Riello, and Barbara Corsetti. Touch-dynamics based behavioural biometrics on mobile devices–a review from a usability and performance perspective. *ACM Computing Surveys (CSUR)*, 53(6):1–36, 2020.

[9] Raffaele Gravina and Qimeng Li. Emotion-relevant activity recognition based on smart cushion using multi-sensor fusion. *Information Fusion*, 48:1–10, 2019.

[10] Jun Han, Shijia Pan, Manal Kumar Sinha, Hae Young Noh, Pei Zhang, and Patrick Tague. Smart home occupant identification via sensor fusion across on-object devices. *ACM Transactions on Sensor Networks (TOSN)*, 14(3-4):1–22, 2018.

[11] Pranjal Kumar and Siddhartha Chauhan. Human activity recognition with deep learning: overview, challenges and possibilities. *CCF Transactions on Pervasive Computing and Interaction*, pages 1–29, 2021.

[12] F. Montalto, C. Guerra, Valentina Bianchi, Ilaria De Munari, and P. Ciampolini. *MuSA: Wearable Multi Sensor Assistant for Human Activity Recognition and Indoor Localization*, volume 11, pages 81–92. 07 2015.

[13] Deepti Sehrawat and Nasib Singh Gill. Smart sensors: Analysis of different types of iot sensors. In *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, pages 523–528, 2019.

[14] Piyush K Sharma, Mark Dennison, and Adrienne Raglin. Iot solutions with multi-sensor fusion and signal-image encoding for secure data transfer and decision making. *arXiv preprint arXiv:2106.01497*, 2021.

[15] Benjamin K Sovacool and Dylan D Furszyfer Del Rio. Smart home technologies in europe: a critical review of concepts, benefits, risks and policies. *Renewable and sustainable energy reviews*, 120:109663, 2020.

[16] Md Shahriar Tasjid and Ahmed Al Marouf. Leveraging smartphone sensors for detecting abnormal gait for smart wearable mobile technologies. *iJIM*, 15(24):167, 2021.

[17] Akriti Verma, Valeh Moghaddam, and Adnan Anwar. Data-driven behavioural biometrics for continuous and adaptive user verification using smartphone and smartwatch. *arXiv preprint arXiv:2110.03149*, 2021.

[18] David H Wolpert. Stacked generalization. *Neural networks*, 5(2):241–259, 1992.

[19] Wencheng Yang, Song Wang, Nor Masri Sahri, Nickson M Karie, Mohiuddin Ahmed, and Craig Valli. Biometrics for internet-of-things security: A review. *Sensors*, 21(18):6163, 2021.

[20] Eric Zeng and Franziska Roesner. Understanding and improving security and privacy in multi-user smart homes: a design exploration and in-home user study. In *28th {USENIX} Security Symposium ({USENIX} Security 19)*, pages 159–176, 2019.