

Leveraging Physical Access Logs to Identify Tailgating: Limitations and Solutions

Carmen Cheh^{*}, Uttam Thakore^{*}, Binbin Chen[†], William G. Temple[‡] and William H. Sanders[§]

^{*}Department of Computer Science, [§]Department of Electrical and Computer Engineering

University of Illinois, Urbana, Illinois, USA

[†]Advanced Digital Sciences Center, Singapore

Email: {cheh2,thakore1,whs}@illinois.edu, {binbin.chen,william.t}@adsc-create.edu.sg

Abstract—Critical infrastructure facilities use physical access systems to control movement in their facilities. However, the cyber logs collected from such systems are not representative of all human movement in real life, including “tailgating”, which is an important problem because it potentially allows unauthorized physical access to critical equipment. In this paper, we identify physical constraints on human movement and use those constraints to motivate several approaches for inferring tailgating from card tap logs. In particular, using our approach, we found 3,999 instances of tailgating in a railway station during a 17-month period. However, certain movement scenarios are not visible in card tap logs. We overcome that limitation by leveraging additional physical data sources to provide information regarding the physical presence of people within a space. We support our findings with an observation experiment that we conducted in a railway station.

Index Terms—physical access system, human movement, tailgating, cyber-physical system, indoor location

I. INTRODUCTION

In critical infrastructure facilities such as power substations, airports, and railway stations, physical security is of the utmost importance, in part because the equipment located within those buildings must be safeguarded from unauthorized insider threats [1], [2]. Physical security is typically enforced using physical access systems that limit human movement in indoor settings. The data collected from such systems are used for a variety of analyses, including detection of unauthorized physical movement. Those analyses rely on the assumption that the cyber data on physical movement reflect actual movement scenarios. However, that assumption is not true in real life.

Physical access systems use door movement sensors and card readers to capture human movement. Those devices can be circumvented by physical and social means, e.g., “tailgating” which means following an authorized person. Tailgating is a serious issue because it could allow unauthorized users into critical spaces, and that can constitute a security violation.

Thus, it is important (1) to determine the extent to which physical access systems can capture tailgating, and (2) to leverage additional data sources to identify the scenarios that cannot be captured by those systems. In this paper, we define physical constraints on human movement that we then use to infer tailgating from data collected by physical access systems. We study the limitations of those cyber data and propose solutions that use physical data sources to complement those

data to identify potential physical security violations. We base our study on a railway system in which a physical access system is deployed to monitor the movement of staff members and visitors in a railway station. The card tap logs collected by that system contain information about entries (card taps) and exits (door movement) to a space. We also collected physical and cyber data from other sources in that railway station, such as sign-in log book entries, to complement the card tap logs. Specifically, our contributions in this paper are as follows:

- We study how tailgating manifests in real life based on observations of tailgating in a railway station over a 6-day period. We use our observations and domain knowledge to infer movement behavior from card tap logs.
- In Section IV, we use the topological constraints on human movement to implement a topology-based checker that identifies tailgating when it finds a discontinuity in a person’s movement trajectory. We thus found 3,999 instances of tailgating in a railway station during a 17-month period.
- In Section V, we use the constraints on space occupancy to develop two approaches, *sum* and *ctr*, based on comparison of the numbers of card taps and egress events in the card tap logs. We show that the *sum* and *ctr* approaches are unable to identify certain scenarios; e.g., a visitor who tailgates in and out of a space remains invisible. Thus, we develop the *ctr-eqp* approach that uses physical data sources such as manual sign-in logs to supply information regarding expected room occupancy, in addition to the card tap logs. We thus found 42 instances of tailgating out of 80 visits to a server room.

II. BACKGROUND AND RELATED WORK

Physical access systems typically use card readers and door movement sensors to provide dashboard information about users’ accesses and door movement [2], [3]. That information can be used to track the locations of users, prevent unauthorized access, and detect violations of access control policy [4], [5], [6], [7]. However, physical access systems do not capture all potential user movement, e.g., tailgating.

Tailgating is a vulnerability that can result in theft, damage of property, and other unauthorized activity that is harmful to the system, and it has been found to occur at a rate of

40 to 60% in some office buildings, according to one report [8]. Measures such as employee education and installation of turnstiles [9] have been taken to prevent tailgating, as recommended by various guidelines, including the NERC CIP-014 [10]. However, physical mechanisms like turnstiles are not enforceable throughout a secure building and it is hard to ensure that all employees will follow rules, even when they have been provided with training.

There has been work on detecting tailgating by using different technologies, such as video surveillance and additional sensors and badges. Advances in computer vision have allowed video surveillance to be further automated to detect and classify motion trajectories [11], [12], [13]. In particular, such algorithms can be used in tandem with radio-frequency identification (RFID) based technologies to detect tailgating [14]. Indoor location systems track users by using networks of beacons on access doors, and RFID tags [15] or phone apps [16], [17], [18] that are located on each user.

The limitation of those technologies is that they require additional installations of physical equipment (like antennas) and distribution of tracking devices, and they also introduce privacy and security concerns. Unlike those approaches, we use existing data from physical access control systems. We view video camera surveillance as a complementary solution and alternative data source to be used as part of our approach to identifying instances of tailgating.

III. CASE STUDY

A railway station consists of a single building that may house one or multiple railway lines through it. Figure 1 depicts the railway station in our case study. The railway station contains 62 rooms that house the equipment necessary to maintain the running of the station and its portion of the railway track. The railway staff can access those spaces only by tapping their access cards at readers on the doors. All the external-facing doors possess card readers, which prevent members of the public from entering prohibited spaces. Although most of the doors inside the staff-only spaces have card readers, there are a number of doors that allow free access.

A. Data Sources

Through a project partnership, we have gained deep knowledge about the physical access system used by the railway system. Below, we describe the data collected by that physical access system, and other supplementary information regarding the presence of people within the station.

1) *Card Tap Log*: We have card tap logs collected by the physical access system. They contain information about card taps and door movement events in a railway station. The events took place between April 2016 and August 2017, and during the month of January 2019. A total of 298,799 card taps were made by 781 users. The logs contain the following information regarding physical movement: (1) timestamp; (2) doorcode; (3) card number; (4) user identification; (5) type of event (**Legal Access** (legal entry), **Invalid Attempt** (failed entry), **Free Egress**); and (6) condition of door (**Door Open Fail**,

Door Close Fail). A failed entry implies either that the user's card has expired or that the user does not have permission to access the room. The condition of the door reflects special situations in which the door has either been left open for more than 10 seconds (**Door Close Fail**), or has not been opened after someone tapped his or her card (**Door Open Fail**).

2) *Equipment Room Sign-in Log*: Each equipment room in the station has a log book. Every person who enters an equipment room needs to record his or her visit in that log book. At the end of their visits, people must record their sign-out times. We collected 143 entries to a server room in the railway station that took place between April 2016 and December 2017. The log book contains the following information: (1) name, (2) the department to which the user belongs, (3) purpose of visit, (4) date of visit, (5) time of sign-in, (6) time of sign-out, and (7) name of accompanying staff member. Only staff members who are custodians of a room have card access to that room.

3) *Manual Observations*: To complement the data that we collected, we conducted a small experiment to elucidate the movement of people in real life as they perform their tasks in the railway station. Given the sensitive nature of the data collection, we observed the movement of people only from public areas such as the concourse and platform. From our vantage point, we gathered a total of 84 movement events through 5 doors from January 7 to 13, 2019 for a total of 8 hours. Those 5 doors serve either as an (1) entrance to the station, or (2) entrance to corridors that lead to equipment rooms, including the server room, as shown in Figure 1a. We recorded the following information: (1) date and time of movement, (2) doorcode, (3) description of people moving, (4) type of movement (entry or exit), and (5) identity of person who taps his or her card.

B. Data Preprocessing

Coalescing. In the card tap log, we found occurrences of multiple **Legal Access** (by the same card) or **Free Egress** events at the same door within the span of a few seconds. Consecutive **Legal Access** events within such a short period are due to a person's tapping of a card more than once to open the door. Similarly, closely spaced **Free Egress** events occur because of multiple movements of the door as someone pushes it open to exit. Since those events are repetitive instances of a single action, we coalesce each of those groups of events into a single entry. After analyzing the distribution of inter-event times, we set the threshold for coalescing at 20 seconds.

Grouping visits. Multiple entries in the equipment room sign-in log represent a single visit by a group of people. Since we want to analyze the movement during each visit separately, we grouped the 143 log entries into incidents representing single visits based on the time period and purpose of the visit. There were a total of 80 incidents, of which 27 were visits by a single person and 53 were visits by a group of people.

Relating observations to logs. We wanted to determine how the movement we observed manifests in card tap logs. Specifically, we wanted to test our hypothesis that if a door

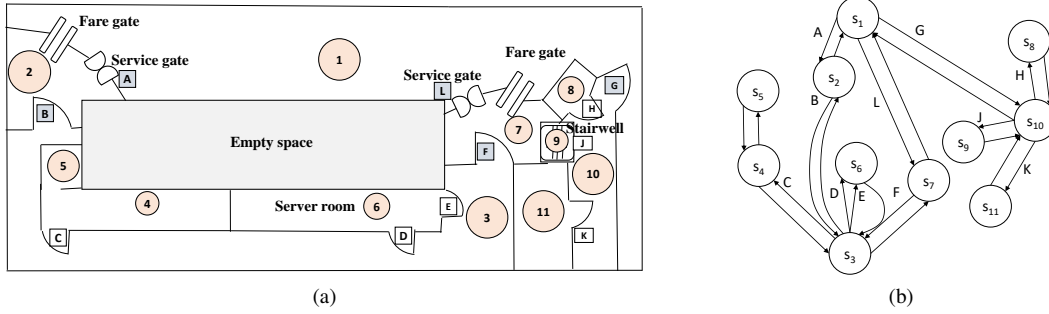


Fig. 1. Building topology of a railway station. (a) A small sample floor plan of Level 1 of the station. The doors that we observe are marked with grey rectangles. (b) Graph representation of (a). Each labeled edge represents a card reader on the door separating the spaces (vertices).

is held open for long, i.e., a **Door Close Fail** event, that implies that tailgating has occurred. So we cross-referenced our observations with the card tap logs by using the recorded timestamps. We found that of the 12 times a **Door Close Fail** event occurred in the card tap logs during the period of our observations, only 3 of those instances corresponded to tailgating. Thus, that disproves our hypothesis and shows that we need a more involved approach to identify tailgating.

C. Physical Constraints on Human Movement

Based on our observations, we found that the events in card tap logs are not indicative of the number of people moving or the direction in which they moved. A **Legal Access** event implies only that the person who tapped his or her card was outside the door and was likely to enter that space. A **Free Egress** event implies that at least one person was inside the space prior to the egress event. So we can only infer the possible locations of people from those events, and thus cannot identify tailgating directly from the card tap logs. Instead, we define two physical constraints on human movement that allow us to use the card tap logs to identify tailgating:

- 1) A person who wants to tap at a card reader *B* in a space that is enclosed by another card reader *A* must first tap at *A* before *B*. A violation of this constraint means that the person tailgated into the space.
- 2) Before a **Legal Access** or **Invalid Attempt** event at card reader (or door) *A*, the occupancy of the space outside the door must be nonzero. Before a **Free Egress** event at door *A*, the occupancy of the space behind the door must be nonzero. So by tracking the occupancy of a space based on card tap log events, we can identify tailgating when there is a violation of the constraint.

We use these two constraints to motivate our approach in Sections IV and V. Our approaches are summarized in Table I.

IV. USING BUILDING TOPOLOGY TO INFER TAILGATING

In this section, we use Constraint 1 defined in Section III-C, that is, a person can only move through a series of connected spaces, to infer tailgating from card tap logs. We build a person's movement sequence by tracking the **Legal Access** and **Invalid Attempt** events. Then, we use the building topology to determine the reachability of the spaces containing the card readers. If there is a gap in the sequence that required the

person to access an additional card reader to reach the space, it implies that tailgating happened there.

More concretely, we represent the building topology as a directed multigraph $G = (S, E)$ in which the set of vertices S represents the spaces in the building. A directed edge $e_i = (v_1, v_2)$ represents possible movement from v_1 to v_2 . The edges are labeled with doorcodes, $label(e)$, if there is a card reader bordering the two spaces. For example, the floor plan in Figure 1a is represented as the graph in Figure 1b.

We use G to determine all possible pairs of doorcodes that can occur in a movement sequence without violating Constraint 1. First, we find all possible simple paths that a person may take between any pair of spaces, $Paths = \{(s_i e_i s_{i+1} \dots s_{i+n} e_{i+n} s_j) | \forall s_i, s_j \in S, e_i = (s_i, s_{i+1}) \in E\}$. Some of the edges along those paths, $Paths_e = \{(e_i \dots e_{i+n}) | (s_i e_i s_{i+1} \dots s_{i+n} e_{i+n} s_j) \in Paths\}$, may not have a doorcode; for example, moving out of a room does not involve tapping a card. So we remove those edges to get $D = \{(e_i \dots e_{i+m}) \subseteq p \in Paths_e | \forall e, \exists label(e)\}$. Then, the possible pairs of doorcodes in a movement sequence are $D_c = \{(label(e_j), label(e_k)) | (e_j, e_k) \subseteq D\}$.

If there is a pair of doorcodes in the movement sequence that does not exist in D_c , $(label(e_i), label(e_j)) \notin D_c$, then the person has tailgated between those spaces, and the set of possible doorcodes that the person skipped is $\{(label(e_{i+1}) \dots label(e_{i+n})) | (e_i, e_{i+1} \dots e_{i+n}, e_j) \in D\}$.

We found a total of 3,999 instances where a person skipped tapping his or her card at a door. 201 out of 781 people had at least one instance of such a violation. An example of a violation was a person who tapped at *L* and then at *D*, skipping *F*. To verify that those instances were indeed an indication that someone skipped a card tap and not an artifact of a disused card reader, we checked that all the possible skipped doorcodes had at least one occurrence of a **Legal Access** recorded in the card tap logs. We found that there always was, and thus, that the 3,999 instances were occurrences of tailgating.

We found that a majority of the missing card taps were to staircases or doors that lead into the station. However, a sizable amount of missing card taps were to corridors that lead to critical equipment rooms (highlighted in a darker shade in Figure 2), although some of those corridors are potentially more tightly controlled and, thus, have fewer missing card taps. So our approach can be used to discover which areas in a building have higher occurrences of tailgating.

TABLE I
COMPARISON OF OUR TOPOLOGICAL-BASED APPROACH AND SPACE OCCUPANCY-BASED APPROACHES.

		Input Data	Target Location	Limitations	# Identified Tailgating
Topology-based approach		Card tap log: Legal Access events	Entire building	Doesn't detect people who tailgates into several spaces and finally a room Doesn't detect people who tailgate into and out of a space Doesn't consider Free Egress events	3,099
Space occupancy-based approaches	<i>sum</i>	Card tap log: Legal Access, Free Egress events	A space	Doesn't detect people who tailgate into and out of a space Doesn't consider sequence of Legal Access and Free Egress events	18
	<i>ctr</i>	Card tap log: Legal Access, Free Egress events	A space	Doesn't detect people who tailgate into and out of a space	20
	<i>ctr-eqp</i>	Card tap log: Legal Access, Free Egress events Sign-in log	An equipment room	Only applicable to equipment rooms	42

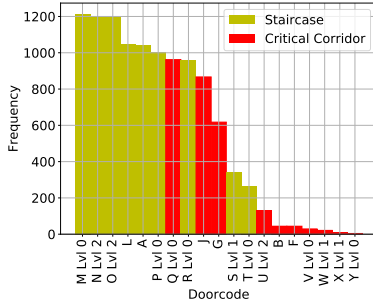


Fig. 2. Frequency of occurrences of violations at different locations.

However, that approach only considers card taps. It does not consider egress events or cases where a person tailgates into several spaces and finally a room (e.g., L , F , D). We will tackle this issue in Section V, where we consider both card tap and egress events.

V. TRACKING ROOM OCCUPANCY TO INFER TAILGATING

In this section, we use Constraint 2 defined in Section III-C, that is, tracking the occupancy of a space, to infer tailgating. However, card tap logs do not provide information about the number of people moving into or out of a space, which is important for understanding the effect of movement on occupancy change. We propose to classify all possible movement scenarios and use our manual observations to identify the most common scenarios for the purpose of assigning movement counts to the events in the card tap logs.

A. Classification of Physical Movement Behavior

The events in the card tap log that pertain to the movement of people are **Legal Access** and **Free Egress**. Since those events are not indicative of the actual movement of people, it is possible for an arbitrary number of people to enter and exit the space. So we define the classes of movement scenarios as the different combinations of counts of people moving **In** and **Out** of the space for each **Legal Access** and **Free Egress** event, i.e., each class is labeled as **Access-In*-Out*** or **Egress-In*-Out***, where * represents a quantifier (**0**, **1**, or **N**) for the number of people moving in the given direction. For the **Access** class, the person associated with the tapping event may or may not move **In** to the space. Independent of that person,

an arbitrary number of people can move **In** or **Out** of the space. So the set of possible quantifiers for the **Access** class is **Access-In{0,1}-Out{0,N}**. For the **Egress** class, the person who pushed the door open may or may not move **Out** of the space. Independent of that person, an arbitrary number of people can move **In** or **Out**. However, unlike the **Access** class, there is no information on who pushed the door open. So we only differentiate between **0**, **1**, and **N** people moving **Out**. So the set of possible quantifiers for the **Egress** class is **Egress-In{0,N}-Out{0,1,N}**. All movement scenario classes are shown in Table II.

We used our manual observations to relate each movement scenario class to real-life situations, which we describe in Table II. We annotated each class with the number of occurrences we found in our observations. We found occurrences of the **Access-In0-OutN** and **Access-In0N-Out0** class that happened at the service gates because those were the only doors that required someone to tap his or her card (**Legal Access**) to exit the space. Therefore, events that occur at the service gates cannot differentiate between a person entering and exiting the station.

By the definition of tailgating, the movement scenarios that involve the movement of a group of people into a space are the only classes that constitute tailgating; we italicize them in Table II. We are particularly concerned with identifying instances of those classes. We can see from Table II that of the movement classes that correspond to tailgating, the **Access-In1N-Out0** and **Egress-InN-Out0** classes happen the most frequently. Thus, we will focus our efforts on identifying instances of those classes. We also find that the most common classes of movement are **Access-In10-Out0** and **Egress-In0-Out1**. So we can assume that a **Legal Access** and **Free Egress** event corresponds to the movement of a single person.

B. Using Card Tap Logs to Identify Tailgating

Based on our domain knowledge, we use the constraint that the occupancy of a space must be 0 at the end of a working day. In tracking the occupancy of a space, from Section V-A, we use the assumption that one event in the card tap logs corresponds to the movement of one person. Then, if no tailgating occurs, we expect that the number of entries will

TABLE II

CLASSES OF MOVEMENT SCENARIOS ANNOTATED WITH THE NUMBER OF OCCURRENCES OF THAT CLASS WE FOUND IN OUR 86 OBSERVATION EVENTS (**FOUND**), AND A DESCRIPTION OF HOW THAT CLASS MANIFESTS IN REAL-LIFE SITUATIONS (**SITUATION**).

Found	Class	Situation
0	Access-In10-OutN	People exit coincidentally as someone taps card
28	Access-In10-Out0	Someone taps card and enters
0	Access-In1N-OutN	Someone taps card and lets people in because of social courtesy or to let visitors in; people exit coincidentally
10	Access-In1N-Out0	Someone taps card and lets people in because of social courtesy or to let visitors in
0	Access-In0N-OutN	Someone taps card for other people to enter but doesn't enter; people exit coincidentally
1	Access-In0N-Out0	Someone taps card for other people to enter but doesn't enter
13	Access-In00-OutN	Someone taps card and no one enters; people exit coincidentally
0	Access-In00-Out0	Someone taps card and no one enters or exits
0	Egress-InN-Out1	Someone pushes door open to exit; people enter coincidentally
0	Egress-InN-OutN	Someone pushes door open for group of people to exit; people enter coincidentally
3	Egress-InN-Out0	Someone pushes door open for people to enter
23	Egress-In0-Out1	Someone pushes door open and exits
6	Egress-In0-OutN	Someone pushes door open for group of people to exit
0	Egress-In0-Out0	Someone pushes door open and no one enters or exits

be equal to the number of exits for that space for each day. So we count the number of entries to a space s_r (**Legal Access** to any door that leads to the space $e = (s, s_r)$) and subtract the number of exits from that space (**Free Egress** from any door that leads to the space $e = (s, s_r)$). If the result is negative, there are more exits than entries, which implies that someone has tailgated into the space.

We define an approach, *sum*, that performs that calculation at the end of the day. We applied *sum* to the server room for the time periods recorded in the equipment room sign-in log so we could compare the results with the analyses later, in Section V-C. The *sum* approach identified 18 occurrences of tailgating. However, the *sum* approach does not take into account the space occupancy during the day. So we defined an approach, *ctr*, that extends *sum* by keeping a running count of the number of entries and exits to the room. Then, by Constraint 2 in Section III-C, we can identify tailgating as soon as the room occupancy falls below 0. The *ctr* approach identified 2 more occurrences of tailgating than *sum* did, and thus performed better than *sum*. However, those two approaches do not take into account people who tailgate both in and out of the room. We tackle this issue in Section V-C.

C. Using Sign-in Logs in Tandem with Card Tap Logs

A railway station consists of staff rooms, store rooms, and equipment rooms. We focus on equipment rooms because they contain critical equipment for the running of the system. Those equipment rooms have sign-in logs that contain information about the number of people in a group during a visit, so we can use that information to determine room occupancy more accurately. Ideally, the number of people in the sign-in log would be equal to the number of people who tapped their cards. However, if a person taps to enter the room but is not in the sign-in log, that is a policy violation, and we indeed found instances of missing names of accompanying staff members in the sign-in logs. On the other hand, if a person appears in the sign-in log but did not tap to enter the room, the *sum* and *ctr* approaches will fail to identify tailgating when there were instances of **Egress-In0-OutN** classes during the visit.

Therefore, we propose the *ctr-eqp* approach that extends the *ctr* approach to account for the number of visitors tailgating

into the room. First, we check that the number of people in the sign-in log is equal to the number of people associated with the **Legal Access** events. If the sign-in log shows more people than the number who tapped their cards, someone must have tailgated into the room. Then, we keep a running count of entries and exits, much like the *ctr* approach, except that on the first **Legal Access** event, the running count is further increased by the number of visitors instead of just 1. The *ctr-eqp* approach identified 22 more occurrences of tailgating than *ctr* did. Thus, the tailgating instances identified by the *ctr-eqp* approach are a superset of the instances identified by the *sum* or *ctr* approach, which shows that inclusion of additional data sources like the sign-in logs help to identify tailgating.

VI. DISCUSSION AND FUTURE WORK

Dataset Limitations. Manual sign-in logs are often not fully accurate or complete. We found that around 60% of the accesses to the server room were not logged as visits in the sign-in logs. The sign-in and sign-out times in the logs are sometimes missing or incomplete. Thus, it is important to understand and acknowledge the limitations of the sign-in log and perform further cross-checks of the information in that log with other forms of data sources.

Approach Generality. Our topology-based approach is applicable to any building since it is based on physical limitations of human movement in an indoor setting. Of our space occupancy-based approaches, *ctr* is also applicable to any building because it capitalizes on the constraint that occupancy of a space should not fall below 0. On the other hand, *sum* and *ctr-eqp* approaches rely on domain knowledge and additional data sources, in the case of *ctr-eqp*.

Attacker Model and Limitations. If an attacker tailgates into any space in the station, they need to have first signed in at the station entrance with a staff member and entered into the space with a non-malicious staff member. We assume that the staff member will ensure that the attacker has logged their entry in the sign-in log book. However, our approaches do not distinguish between “benign” tailgating and “suspicious” tailgating that is the result of malicious behavior. As we see in our results, a large portion of the tailgating instances that we uncovered were due to visitors who needed to be escorted by

railway staff members. In those cases, tailgating was necessary for them to do their jobs. While it can be argued that such instances need not be identified, we strongly believe that it is essential to uncover them, because they give insight into potential policy violations (e.g., a visitor’s borrowing of a staff member’s card for convenience).

Future Work. We can extend our approaches that only look at events in a specific space to consider a person’s taps at other locations. If a person taps somewhere outside a room, we can infer that he has left the room, but that only allows us to associate egresses with that person. We can, however, use other logs that provide us with rich data that can be used to determine human presence and provide context as to whether their presence is for malicious intent. There are other logs that we do not draw on in this paper, i.e., device event logs that describe state changes of devices in equipment rooms, and station sign-in logs that all visitors and staff members need to sign before entering the station. We can use device event logs to determine when people are interacting with the devices in a room. The main challenge with using that data source is that of correctly attributing changes in the device state to physical interaction with a person instead of automated physical processes or remotely controlled actions. Then, we can use that information to establish the presence of people within the room. We can also use the station sign-in logs to identify the people who are in the station at a given point in time. Much like the equipment room sign-in logs, that knowledge can help us find instances of tailgating. However, it is harder to use station sign-in logs than equipment room sign-in logs because we only have information that a certain person is in the building at a given time; we cannot pinpoint his or her location. Thus, we propose to use building topology to track the movement of all users in the building to correlate the **Legal Access** and **Free Egress** events.

VII. CONCLUSION

Physical access systems that rely on card readers and door movement sensors are inadequate to prevent tailgating behavior that potentially violates physical access control policies. We defined several physical constraints on human movement that drove our approach to inferring tailgating from card tap logs. We analyzed the movement trajectories derived from the card tap logs and found 3,999 instances of tailgating throughout the building. Then, we narrowed our focus to critical equipment rooms and proposed two approaches that keep track of room occupancy using only the card tap logs, and another approach that uses equipment room sign-in logs together with the card tap logs to identify tailgating. Our results show that the equipment room sign-in logs are very useful in identifying tailgating, and we discussed ways of using additional data sources to supplement our approach.

ACKNOWLEDGEMENTS

This work was supported by the Maryland Procurement Office under Contract No. H98230-18-D-0007, and in part by the National Research Foundation (NRF), Prime Minister’s Office,

Singapore, under its National Cybersecurity R&D Programme (Award No. NRF2014NCR-NCR001-31) and administered by the National Cybersecurity R&D Directorate. We thank the experts from SMRT Trains LTD for providing us with data and domain knowledge.

REFERENCES

- [1] M. E. Luallen, “Managing insiders in utility control environments,” SANS Institute, Tech. Rep., 2011. [Online]. Available: <https://www.sans.org/reading-room/whitepapers/analyst/managing-insiders-utility-control-environments-34960>
- [2] “AlertEnterprise Guardian Physical,” <https://www.alertenterprise.com/products-EnterpriseGuardianPIAMandBadging.php>, accessed: 2019-03-20.
- [3] “IDenticard PremiSys Security Management Dashboard,” <https://www.identicard.com/access-control/premisys-security-management-dashboard/>, accessed: 2019-03-20.
- [4] C. Cheh, B. Chen, W. G. Temple, and W. H. Sanders, “Data-driven model-based detection of malicious insiders via physical access logs,” in *Proc. 14th International Conference on Quantitative Evaluation of Systems*, N. Bertrand and L. Bortolussi, Eds. Springer International Publishing, 2017, pp. 275–291.
- [5] A. Gellert and L. Vintan, “Person movement prediction using hidden Markov models,” *Studies in Informatics and Control*, vol. 15, no. 1, pp. 17–30, 2006.
- [6] C. Koehler, N. Banovic, I. Oakley, J. Mankoff, and A. K. Dey, “Indoor-ALPS: An adaptive indoor location prediction system,” in *Proc. ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2014, pp. 171–181.
- [7] J. P. Boyer, K. Tan, and C. A. Gunter, “Privacy sensitive location information systems in smart buildings,” in *Proc. of the 3rd International Conference of Security in Pervasive Computing*. Springer, 2006, pp. 149–164.
- [8] AlliedUniversal Security Services, “Security tailgating (aka piggybacking),” Security, Resiliency and Technology Integration Forum, Tech. Rep. [Online]. Available: <http://www.alliedbarton.com/Portals/0/SRC/WhitePapers/Security%20Tailgating%20-%20Best%20Practices%20in%20Access%20Control.pdf>
- [9] L. Fennelly and M. Perry, *Physical Security: 150 Things You Should Know*. Butterworth-Heinemann, 2017.
- [10] C. Vinson, J. Hallenstein, R. L. Fisher, and S. D. Perusquia, “Review of physical security protection of utility substations and control centers,” State of Florida Public Service Commission Office of Auditing and Performance Analysis, Tech. Rep. [Online]. Available: http://www.psc.state.fl.us/Files/PDF/Publications/Reports/General/Electricgas/Physical_Security_2014.pdf
- [11] T. K. Ho, K. Matthews, L. O’Gorman, and H. Steck, “Public space behavior modeling with video and sensor analytics,” *Bell Labs Technical Journal*, vol. 16, no. 4, pp. 203–217, March 2012.
- [12] H. Liu, S. Chen, and N. Kubota, “Intelligent video systems and analytics: A survey,” *IEEE Transactions on Industrial Informatics*, vol. 9, no. 3, pp. 1222–1233, Aug. 2013.
- [13] K. F. MacDorman, H. Nobuta, S. Koizumi, and H. Ishiguro, “Memory-based attention control for activity recognition at a subway station,” *IEEE MultiMedia*, vol. 14, no. 2, pp. 38–49, April 2007.
- [14] R.-S. Hsiao, T.-X. Chen, C.-H. Kao, H.-P. Lin, and D.-B. Lin, “An intelligent access control system based on passive radio-frequency identification,” *Sensors and Materials*, vol. 29, no. 4, pp. 355–362, 2017.
- [15] J. Toledo-Castro, P. Caballero-Gil, N. Rodríguez-Pérez, I. Santos-González, and C. Hernández-Goya, “Beacon-based fuzzy indoor tracking at airports,” in *Multidisciplinary Digital Publishing Institute Proceedings*, vol. 2, no. 19, 2018, pp. 1255:1–1255:9.
- [16] “AnyPlace,” <http://anyplace.cs.ucy.ac.cy/>, accessed: 2018-05-20.
- [17] “infSOFT,” <https://www.infsoft.com/industries/airports/features>, accessed: 2018-05-20.
- [18] J. Xiong and K. Jamieson, “Arraytrack: A fine-grained indoor location system,” in *Proceedings of the 10th USENIX Conference on Networked Systems Design and Implementation*, 2013, pp. 71–84.