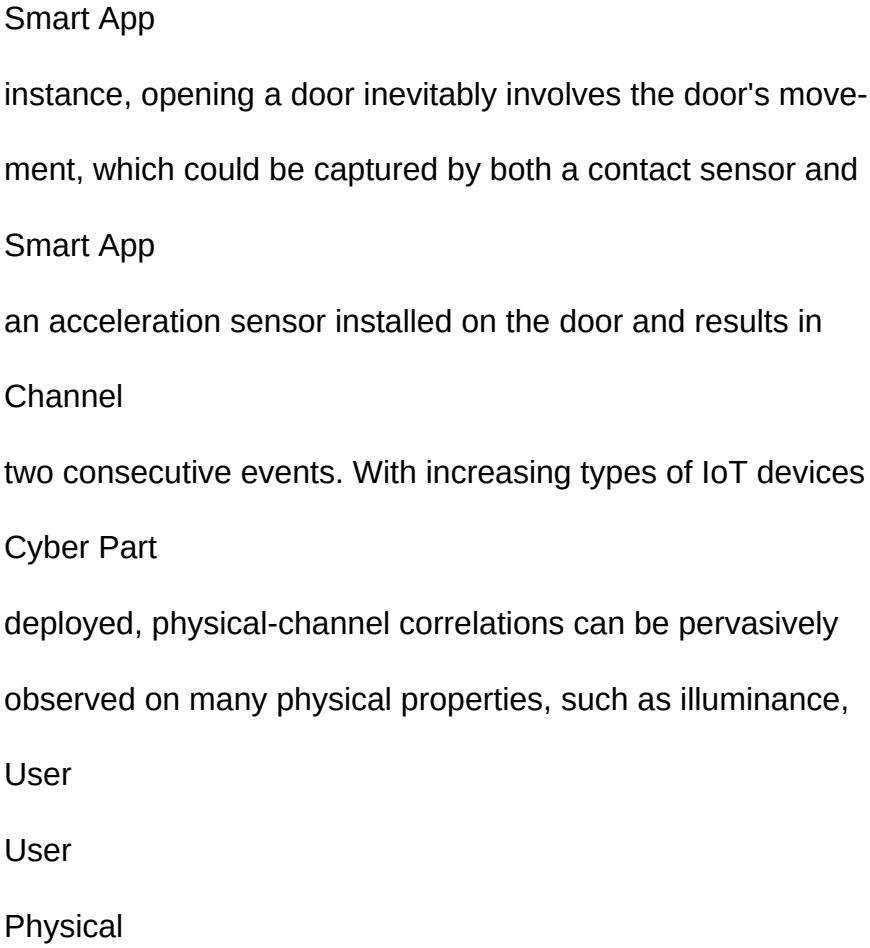

Figure 3: Correlation channels.

instance, opening a door inevitably involves the door's movement, which could be captured by both a contact sensor and an acceleration sensor installed on the door and results in two consecutive events. With increasing types of IoT devices deployed, physical-channel correlations can be pervasively observed on many physical properties, such as illuminance, power, sound, and temperature [39].

User Activity Channel. While user activities impose changes on devices, device states also reflect user activities. Thus, the user activity channel causes correlations between

We assume the IoT platform is not compromised. Like
devices. For example, a TV being turned on typically implies
other anomaly detection work [35,51,76], we assume there
that the user is nearby, which should be captured by the
are no or very few anomalies during training. We assume
motion sensor. When a user returns home, there should be
there are no malicious or conflicting rules in the installed
consecutive events, such as "presence on" showing the user's
smart apps; how to detect malicious logic [71] and conflicting
proximity and "contact-sensor open" for door opening.
rules [28,34] are two separate research problems, and there
4.2 Representation of Correlations
are existing solutions to them [28,71], including our prior
work [33,34]. Gartner predicts that a typical household could
An event reporting that the device A's attribute a should
have more than 500 IoT devices by 2022 [72]. Given the
be changed to the value a is denoted as (a(A) while a
state
dense deployment in the near future, we exploit scenarios
which indicates that the device B's attribute has the value
where an IoT device has one or more other devices nearby
b is denoted as B(B) 2 We define two types of correlations.
to interact with, and propose to leverage them to detect
a device's anomalous physical behaviors. We discuss the
The event-to-event (e2e) correlation. It means that one
case of no interactive devices nearby in Section 8. Jamming

event should be followed by (denoted as another. For

that blocks communications reporting IoT events can be

example, given a motion sensor A and a light B, the e2e

motion(A)

easily detected due to session timeout or missing sequence

correlation 'active

Eon

.switch(B)

means

the

event

numbers; we thus do not further discuss it.

_motion(A)

'active

should be followed by the event Eon -switch(B)

4 Correlations

The event-to-state (e2s) correlation. It means that

Devices deployed in the same home may correlate in the

one event arising implies (denoted as my) a state

form of co-present or temporally related events [35,39,45,68].

is true. For example, power(plug) switch(heater)

These correlations can be attributed to the execution of smart

means that, when the arises, the state

apps [29], physical interactions [39] or users' activities [45].

switch

As shown in Figure 3, we investigate the causes of these

Son

(heater)

should

be

true.

correlations and categorize them into three channels below.

For the representation of a correlation involving condi-

4.1

Correlation Channels

tions, its anterior event is combined with the conditions

using the "A" symbol. For example, ^ Presence present

Smart App Channel. Smart apps not only directly cause

Eon

.switch(Light)

means the event Motion active if the condition

correlations between triggers and actions as programmed,

present Presence is true, should be followed by Eon ,switch(Light)

but also imply some extra correlations that should be consid-

We show in Section 5 that the two types of correlations,

ered. For example, the smart App "light follows me" [2] leads

despite their simplicity, are very effective in capturing rich

to the correlation between the motion sensor and the light,

semantic information and modeling the relations of devices

and also implies a possible correlation worth verification,

that correlate via different channels.

that is, "if the light is turned on, then the motion should be in

the active state". The implied correlation is true if the light is

5

HAWatcher Design and Implementation

exclusively turned on by the smart app.

We first introduce the workflow of anomaly detection (Sec-

Physical Channel. Two devices can correlate via a cer-

tain physical property. First, an actuator device's action can

tion 5.1), and then describe the major modules in HAWatcher,

as shown in Figure 4: 1) Semantic Analysis (Section 5.2), 2)

change a physical property, which is captured by nearby

sensor devices observing that property. For example, a smart

Correlation Mining (Section 5.3), 3) Correlation Refining

light's action can affect an illuminance sensor nearby. Second,

(Section 5.4), and 4) Anomaly Detection (Section 5.5).

different sensor devices can be affected by the same physi-

For simplicity of description, without causing confusion we sometimes

cal event and generate temporally correlated IoT events. For

omit the device IDs and use the simplified notations E and sB.