6thSense: A Context-aware Sensor-based Attack Detector for Smart Devices

Sowmya Jayaram Iyer
jayarami@purdue.edu
PUID: 0033742039

SUMMARY:

The paper presents a novel design for a comprehensive context-aware intrusion detection system. The two main contributions mentioned were: (1) 6thSense achieved over 95% accuracy in classifying malicious behavior from user incidents in real-time utilizing three different machine learning models (Markov Chain, Naive Bayes, and LMT) and (2) 6thSense yields minimal overhead. Previous works mentioned (Audroid, Semadroid and Darkly) focused mainly on a single or similar type of attack scenarios or sensor threats while in real scenarios attackers can exploit the sensors in numerous other ways utilizing multiple sensors. This paper thus includes three threat models: using sensors to (1) Trigger malicious apps (2) leak sensitive information of users (3) steal information through inference from device modes) and presents an adversary model to tackle the same. The authors tested the proposed framework and presented the experimental results on data collected from 50 real users against the aforementioned threats. In this evaluation, the authors show an accuracy of above 96% to prove that the approach is effective and can be improved. The approach seems to be novel with limited performance overhead useful for practical applications, however, the paper seems to be unclear along several lines such as data collection, for instance, making the technical results appear unreliable. The author has done a good job in conveying the need for this approach, however, needs several improvements in showing the reliability of the data and the robustness of the solution. Comments and suggestion are listed in the next section

DETAILED COMMENTS:

Existing sensor management systems, specifically Android sensor management system, in Smart devices heavily rely on permission-based access control. This permission is asked to the user only while installing an app and, once granted, user loses control over the usage of these listed as well as unlisted sensors by the App. The concept of not seeing sensors as an individual entity but as a set of entities specific to each task performed by a user while considering sensor-based threats is a new yet conceivable idea. However, 6thSense depends on the fact that the model is aware of the sensors activated by each and every user task or activity. This, however, is not elaborated upon or included while

describing the data and only 9 activities are considered and tested upon. In the line “Our evaluation shows that 6thSense can detect sensor- based attacks with an accuracy and F-Score over 96%” does not mention which model was chosen in the final model or was an ensemble used.

One of the drawbacks in the previous works mentioned was in Darkflow which simply goes over the trust levels of different applications. This work does not provide a real-time threat detection solution and thus, the real-time nature of 6thSense addressing this problem by moving from App-based threat detection to Sensor-based detection system is more reliable. Also, 6thSense appears to give a robust solution for threats across multiple sensors while Audroid and Semadroid fail to do so. Thus, this solution does a good job addressing real-time threats faced by smart device users

The proposed model is based on four basic assumptions. Firstly, the sensors are not considered as an individual entity but as a set of “co-dependent” entities specific to an activity or task. Secondly, the data is sampled over a particular period of time. Thirdly, Machine Learning algorithms provide a faster technique to process the large amount of data from different sensors. Lastly, real-time monitoring mitigates the possibility of data tampering or false data injection.

6thSense uses three main ML techniques, Markov Chain, Naïve Bayes and LMT for the analysis of sensor data. To relieve dependency on computational resources for the large amount of sensory data, Markov Chain is used. Here, the prediction is don’t on a transition matrix which determines the probability of transition between two states at a given time. The main assumption of the Naive Bayes detection is that the presence of a particular sensor condition in a task/activity has no influence over the presence of any other feature on that particular event. However, why this was used against/with Markov Chain is not specified. Similarly, the choices for alternate ML models seems rather random and is not well elaborated upon. Also. author mentions that the dataset consists of specific user activities and lab-induced malware activity, discussions on how a Logistic Model Tree would differentiate between a new accepted activity and new malware and what are the consequences need to be included.

6thSense framework has three components: Data Collection phase, Data Processing phase and the data Analysis phase. All the three phases require improvements along the following lines:

Data Collection Phase:

While the author mentions the importance of sensor-based threat detection methods for all sensors, he has chosen nine sensors available and deemed the others not influenced by all typical user activities hence unnecessary to be included. However only 9 activities with around 3-5 instances of 5 minutes length each are observed. This data is clearly not populated enough to validate the 96% accuracy as global accuracy.

Data Processing Phase:

Figure 2 shows 5 steps in data preprocessing phase starting from removing wrong data, sampling, determining sensor state, merging data from other Applications, building data metrics. However, only the sampling and assigning sensor state steps are elaborated or at the least mentioned.

The following three threats are considered by 6thSense: (1) a malicious App that can be triggered via light or motion sensors, (2) a malicious App that can leak information via audio sensor, and (3) a malicious App that steals data via camera. Under the heading Training Environment, elaboration on the data is given rather than about training environment itself and hence a better heading could be used to not hamper the flow of the reader. Here, the author describes how the malicious dataset was created. In Page 10, "For Threat 1, we developed two different Android Apps which could be triggered using the light sensor and motion sensors on the smartphone. To perform the attack described in Threat 2, we developed a malware …. For Threat 3, we developed a malicious App …". Author repeatedly mentions developing a malware. The malicious apps created for the dataset, are they developed totally from scratch (they may fall under new malware then, not found in any preexisting database), or are they based on other pre-existing malicious software (and if so then what).

The evaluation of Markov Chain-Based Detection showed a 98% accuracy while setting a threshold of 3 above which the number of consecutive malicious states on the device denotes a malicious app. However, this threshold is clearly hardcoded and hence require a huge amount and variety of malware attacks to be used as a universal cut-off threshold. The evaluation of Naïve Bayes showed an accuracy of 95% for a threshold of 60%. This "threshold to determine the correct activity" introduced in Page 12 is not explained properly and hence the results don't give a clear idea of what is happening. We see a clear trade off in false positive rate between 65% and 80% threshold. Author has not explained why 65% threshold was chosen over the 85% threshold with a lower FP rate. As a reader I find minimizing FP more important over accuracy, and hence, the author should substantiate his choice with valid arguments. In alternate ML models, LMT sure does give a very low False Positive rate compared to other algorithms, but the intuition behind it is not elaborated upon. Also, the architecture and working of LMT is nowhere mentioned.

Even while comparing the above three models, the author does not mention how these results were used in the 6thsense framework. LMT performs better than others but if 6thSense is built upon LMT, or was an ensemble model used is not elaborated upon. The results were not discussed in terms of the application as to what a low or high False Positive rate change the outcome in real scenario. The comparison para is too vague and gives a simple qualitative comparison but no information on what needs to be inferred from it.

The author, however, has neatly explained and tabulated the performance overhead of this model which provides sufficient evidence for the feasibility of this solution.

It has been already clearly highlighted that the major limitation in this paper is the validity of the dataset used. Since this data is collected in specific scenarios, its performance in real-life settings is indeed reasonably doubtful. The author has acknowledged that the data uses no real sensor-based malware and that they are lab-made malware. However, to give a comparison with real malware, the author uploads the malware in a software scanner named *VirusTotal*. The observation is that only 2 out of 60 of the malware scanners reported malicious behaviors and the author concludes that the existing security schemes couldn't detect the sensor-based threats. However, without any information on the tool's efficiency for detecting zero-day attacks, since the data is "developed" by the author, this comparison or claim does not hold well.

RECOMMENDATIONS:

This paper does a good job in explaining the need for the model and how it works against different threat models is clearly explained. However, the data is not populated enough to prove the accuracy of the model. Using a peer-reviewed data and also, more information on the lab-made malware (mentioned in the comments) would improve the reliability of the proposed model. The author needs to also provide information on how the real-time monitoring is performed as users still have to use phone sensors on a day-to-day basis without triggering 6thsense, and if so, then malicious apps can easily be installed with or without the knowledge of 6thsense which can tamper with data without necessarily utilizing any EXTERNAL sensors (Ransomware links in email, etc.). At its present form, with improvements suggested in the previous section, the paper makes an acceptable case for publication.