Authors' copy downloaded from: https://sprite.utsa.edu/

Copyright may be reserved by the publisher.

Anindya Maiti *University of Oklahoma, USA*
Murtuza Jadliwala *University of Texas at San Antonio, USA*

Editors: Nic Lane and Xia Zhou

SMART LIGHT-BASED INFORMATION LEAKAGE ATTACKS

Excerpted from "Light Ears: Information Leakage via Smart Lights," from *Proceedings of the ACM on Interactive Mobile, Wearable and Ubiquitous Technologies*, 2019, with permission. © https://dl.acm.org/doi/10.1145/3351256 ACM 2019

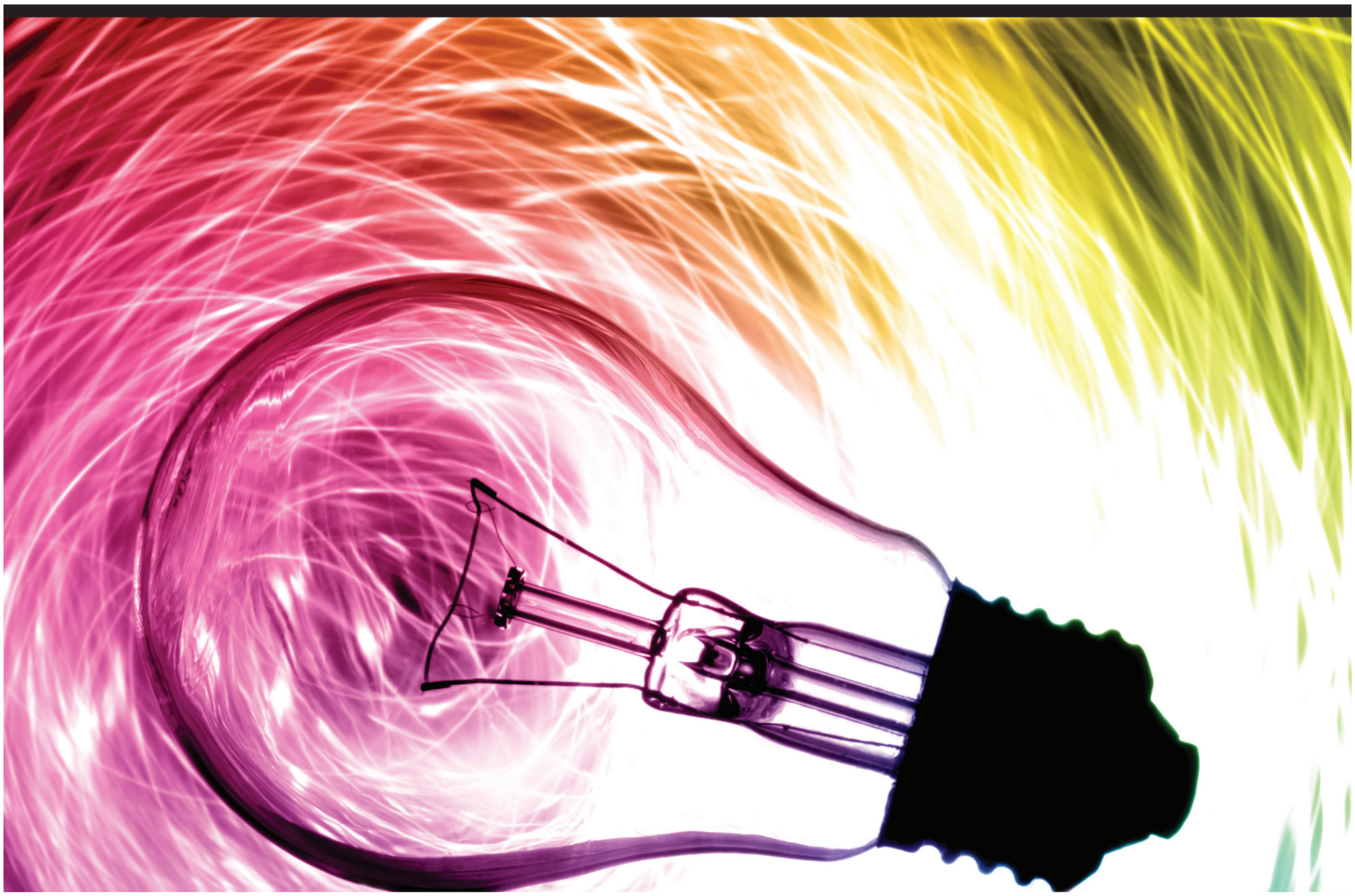
Modern Internet-enabled smart lights promise energy efficiency and many additional capabilities over traditional bulbs. However, these connected lights also expose a new attack surface, which can be maliciously used to violate users' privacy and security. We design and evaluate novel inference attacks that take advantage of the light emitted by these smart lights to infer sensitive user data and preferences.

A popular feature of modern smart lighting systems is the ability to remotely control its functionality over a Wi-Fi, Bluetooth, or ZigBee network. Many current generation smart lights (such as LIFX and Philips Hue) are also LED-based, which enables fine-grained customization of color and intensity of the emitted light. Some advanced smart lights (such as LIFX+) are also equipped with infrared capabilities. Given these new capabilities, this article aims to highlight the susceptible state of personal information of smart light users by outlining specific system vulnerabilities and privacy threats that could take advantage of these vulnerabilities.

We first focus on exploiting a new feature of modern smart lights, known as *multimedia-visualization* (Sections 1 and 2). Multimedia-visualization is intended for use in conjunction with a song or video playing on a nearby or connected media player, which results in a vibrant lighting effect that is synchronized with the tones present in the audio or the dominant colors in the video stream, respectively. However, consider a scenario where a curious adversary is able to observe the changing light intensities/colors of a multimedia-visualizing smart light installed inside a user's residence (likely through a window). *Can the adversary determine what song/video is being played by only analyzing the changing light intensities/colors of the smart light?* Such attacks, if successful, can have significant privacy implications for smart light users.

Further, we also study the feasibility of an adversary exploiting a smart light's infrared functionality to invisibly exfiltrate a user's private data out of his/her secured personal

Illustration, istockphoto.com



device or network (Section 3). We show that such an attack can be accomplished by controlling and carefully manipulating the light emitted by the infrared bulb on these systems to create a "covert-channel" to the adversary on which sensitive data can be exfiltrated.

ADVERSARY MODEL

For the inference threats on the audio-visualizing and video-visualizing functionalities of smart lights, we assume a passive adversary whose goal is to infer a target user's media consumption by visually eavesdropping on the light emitted by their smart bulb, without actively attacking the user's wireless network or appliances. The user's wireless network is assumed to be secured against eavesdropping attacks, for example using WPA2, so the adversary cannot perform direct analysis of the packets sent to the smart bulb. For the data exfiltration threat using infrared-enabled smart lights, we assume an adversary whose goal is to exfiltrate data out of a target user's network or personal device. For the data exfiltration attack to work, the adversary has to additionally install a malicious software agent on the target user's device (for example, their smartphone or computer) that connects to the same network as their infrared-capable smart bulbs. This can be achieved by social engineering attacks, or by tricking the user in to installing a Trojan application. The malicious software agent is responsible for encoding target user's private data (accessible on-device or on the network) in a format suitable for infrared communication, and transmits the encoded data using the user's infrared-enabled smart light.

AUDIO INFERENCE THREAT

In the *audio-visualization* mode, a smart bulb reacts to the high and low tones present in the input audio stream by fluctuating its output light brightness. It became evident from our exploratory experimentation that there exists a clear correlation between the audio waveform of the song and its corresponding "luminance-profile" recorded using a luminance meter. Another observation

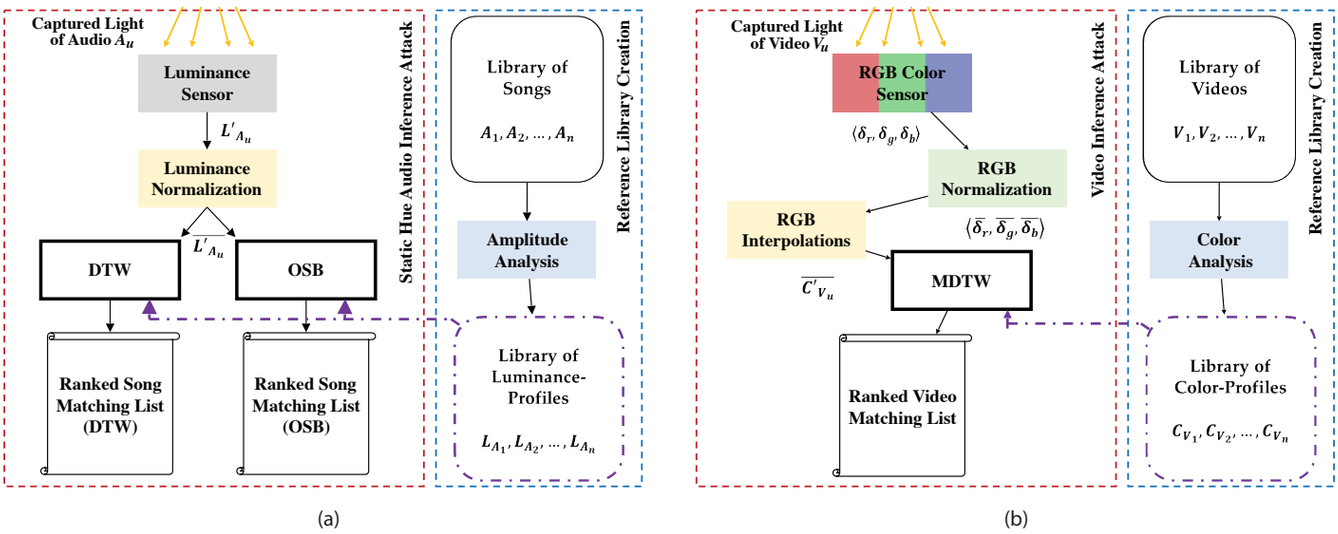


FIGURE 1. (a) Audio inference framework; (b) Video inference framework.

was that for a given song, although the luminance-profile suffers minor distortions across multiple recordings, the similarity between them is generally very high. We utilize these properties to design our audio inference framework, which is based on elastic time-series matching [5] and is immune to minor distortions.

AUDIO INFERENCE FRAMEWORK

Based on the above understanding of audio-visualization properties, we can design an inference framework for inferring a source audio from its corresponding luminance-profile as follows (Figure 1a):

Capturing Luminance-Profile. The inference attack starts with the adversary recording the luminance-profile of an unknown target song using a luminance meter (such as Yoctopuce V3). The observed luminance-profile is a time-series of the observed luminance values.

Luminance Normalization. Once the entire luminance-profile is recorded for a chosen observation duration, the next step is to normalize it to achieve amplitude invariance, which aids in the similarity search later [2].

Creating a Reference Library. Before matching the normalized luminance-profile against a library (of songs), the adversary has to create a reference library of luminance-profiles corresponding to a comprehensive set of songs. Template luminance-profiles can be created by sampling the amplitudes in waveform audio files, and converting them to absolute values. These template luminance-profiles serve as an approximate representation of how an audio-visualizing smart bulb will react.

Similarity Search. The final step for the adversary is to match the observed luminance-profile to songs in the reference library. A classification method based on algorithms that measure similarity between temporal sequences, which are misaligned and vary in time or speed, for example, Dynamic Time Warping (DTW) [5] and Optimal Subsequence Bijection (OSB) [3], can be employed for this purpose. The framework outputs the song (as a prediction), whose template yields the minimal distance between the observed luminance-profile and template luminance-profiles in the reference library.

AUDIO INFERENCE EVALUATION

We complied a reference library of 400 chart-topping songs, which is analogous to password cracking using a dictionary of most commonly used passwords. We test in an outdoor setting (Figure 2) where the observation point was at 50 meters away from the bulb and a 80 *mm* 45-225x telescope was used to focus observed light on a luminance meter. For 100 test songs, we measured the accuracy of the framework based on the rank of predicted songs matched against the entire reference library (rank of 1 being the correct prediction). Some of the key observations were:

- The mean predicted song rank was as low as 1.20 when the adversary is able to record the target song-visualization for 120 sec, which implies very high inference accuracy.
- There exists a higher confusion within the same genre of songs, which implies that even if an adversary is unable to match accurately the audio-visualized data to its corresponding song, the adversary can still infer the user's media or genre preference.

VIDEO INFERENCE THREAT

The *video-visualization* feature enables a smart bulb to react to the colors present in the input video stream by changing its output light color to the average RGB composition of the current frame in the video. Our preliminary observations brought us to two similar conclusions as before. First, the observed RGB color from the bulb has some correlation with the average RGB color in the current frame of the video stream, which is expected. As a result, the RGB "color-profile" of a video observed on a video-visualizing bulb is unique to the video, and the probability that a completely different video also has the same color-profile is small. Second, similar to luminance-profiles in audio-visualization, even though color-profiles suffers minor distortions across multiple recordings, the similarity between them is generally very high.

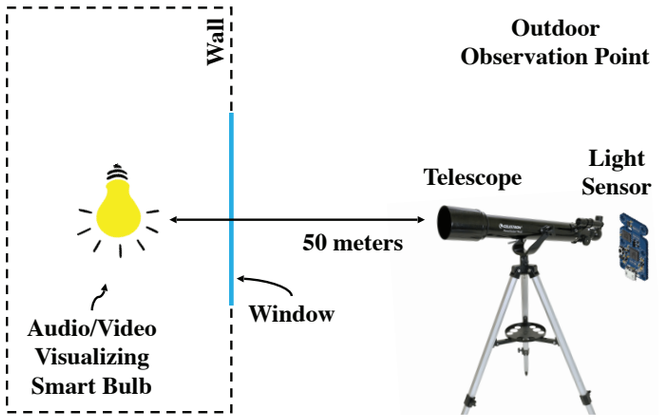


FIGURE 2. Experimental setup for evaluating proposed attacks.

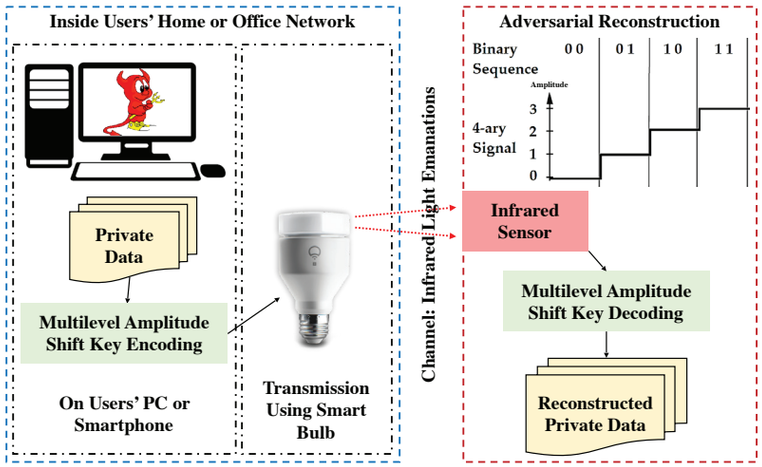


FIGURE 3. Covert data exfiltration framework.

Video Inference Framework

With the above understanding of video-visualizing light properties, we can now design a video inference framework (Figure 1b).

Capturing Color-Profile. The inference attack starts with the adversary recording the observed color-profile of an unknown target video using some RGB color sensor, such as Vernier GDX-LC. The observed color-profile is the time-series of observed RGB values, recorded at a constant sampling rate.

Color Normalization and Interpolation. Once the entire color-profile is recorded for a chosen observation duration, the adversary next normalizes and interpolates it to create the corresponding normalized and interpolated color-profile. The normalized and interpolated color-profile is amplitude invariant and accounts for imperfect brightness scaling, which aids in similarity search.

Creating a Reference Library. The adversary next creates a library of template color-profiles corresponding to video files in a reference library by sampling the RGB composition in these video files. These template color-profiles serve as an approximate representation of how a video-visualizing smart bulb will react.

Similarity Search. The final step for the adversary is to match the color-profile against the template color-profiles corresponding to the reference library of videos using a time-series similarity computing technique such as *Multidimensional Dynamic Time Warping (MDTW)* [7]. MDTW is a generalization of DTW for measuring similarity between temporal sequences, in two or more dimensions. We compute the 3DTW distance between the observed color-profile and template color-profiles in the reference library, selecting the video (as a prediction) whose template yields the minimal distance.

Video Inference Evaluation

We complied a reference library of 500 full-length movies released on DVD and Blu-ray in the last 10 years. We test from the same outdoor observation point as in the audio inference evaluation (Figure 2). Some of the key observations are:

- The mean predicted video rank was as low as 1.49 when the adversary is able to record the target video-visualization for 360 sec, which implies very high inference accuracy.
- Our attack framework is functional even when the adversary does not have direct Line-of-Sight to the smart bulb. In Non-Line-of-Sight (NLOS) inference we observed an average difference of +0.17 in mean rank, which is marginally less accurate than LOS inference.

COVERT DATA EXFILTRATION THREAT

Our final attack framework (Figure 3) enables an adversary to employ a smart light's infrared capability to covertly exfiltrate private data out of a user's personal device or network. What makes this attack interesting is that traditional light bulbs are normally not perceived (or monitored) as an attack surface, even in high security establishments. This attack methodology not only shows that a smart light can be used to transmit data, but it also shows how such an attack when carried out from within a secure air-gapped network can become a significant privacy and security threat. Moreover, unlike Internet gateways which can be protected against data exfiltration attacks using a firewall, an exfiltration gateway made out of a smart bulb has no such restrictions.

The Covert Channel. We utilize a multilevel (*M*-ary) amplitude shift keying (ASK) [1] encoding technique to implement an infrared-based covert channel, as our experimental bulb (LIFX+) supports only 950 *nm* infrared light. As human eyes are

Original Text: `A cup of sugar makes sweet fudge`
Reconstructed Text: `A buq pf!sugbr m`kes▮suees hudfe`

FIGURE 4. Text reconstruction using 256-ary ASK at 15 m.

not sensitive to the infrared spectrum, it can be used to create a covert channel, which can remain visually undetected.

Private Data Encoding and Transmission. Once the value of M (number of levels in M-ary ASK) is decided by the adversary based on the distance and acceptable level of reconstruction quality, the next step is to encode private data of interest (in binary form) using M-ary ASK. A malicious software agent installed on the target user's device or network undertakes this task. The encoded data is then transmitted in blocks by controlling the infrared power level of the smart bulb connected to the same device or network.

Adversarial Reconstruction. On the adversary's side, he/she observes the target user's smart bulb using an infrared sensor (such as TSOP48). Once a start symbol is received by the infrared sensor, the adversary starts recording the observed infrared amplitudes representing the M-ary ASK encoded data, until an end symbol is received. Then the adversary normalizes the recorded data based on the maximum amplitude, and decodes it to reconstruct the private data in binary format.

Data Exfiltration Evaluation

The infrared signal strength reduces as distance from the bulb increases. As a result, the boundaries between the M amplitude levels present in a M-ary ASK signal is also diminished, leading to higher confusion between neighboring amplitude levels and thus errors in the reconstructed data. This phenomenon was evident in our evaluation results (Figure 4), especially for higher values of M. The adversary can potentially improve reconstruction by employing Forward Error Correction (FEC), a digital signal processing technique used to enhance data reliability [6]. Nonetheless, the adversary can still extract some useful information without using FEC. Figure 4 shows the example of a reconstructed sentence where the incorrectly reconstructed letters are neighboring to the original letters on the ASCII chart. An adversary can perform simple syntactical and semantical analysis on the reconstructed text to improve its correctness and legibility.

DISCUSSION

Limitations. Setting up a secure observation point in the neighborhood of the target user is a prerequisite towards carrying out the proposed attacks. While it may be difficult for an adversary to maintain covertness in rural areas with fewer structures where he/she can hide, it may be easier to find a secretive observation point in an urban setting. Several external factors can also disrupt or impair observation, which the adversary must account for. For example, light from passing-by automobiles can introduce temporal noise, moving bodies near the target user's window can change light characteristics, and rainy weather can introduce a high degree of unpredictable noise in the observation channel.

Countermeasures. A simple mitigation would be to cover the windows with opaque curtains and block light leakage to the outside. Additionally, for the inference attacks, the maximum brightness of the bulbs can be reduced, so that the light leakage is also reduced. To prevent the exfiltration attack, strong network rules can be enforced such that unauthorized computers and smartphones cannot control smart bulbs over an IP network.

CONCLUSION

We designed and evaluated multiple sensitive information leakage frameworks that exploit modern smart lights. These threats affirm the need for better protection mechanisms, such as, strong access control within smart light management protocols. An extended version of this article is published in ACM IMWUT [4]. ■

Acknowledgments

This work was partially supported by the NSF under award #1943351.

Anindya Maiti is currently an Assistant Professor in the Department of Computer Science at the University of Oklahoma, USA. He received his PhD in Electrical Engineering and Computer Science and his MS degree in Electrical Engineering from Wichita State University, USA, in 2018 and 2014, respectively. His current research interests include vulnerability discovery and remediation in cyber-physical systems, and applied machine learning research in security and privacy.

Murtuza Jadliwala is an Assistant Professor in the Department of Computer Science at the University of Texas at San Antonio, USA. He received his bachelor's degree in Computer Engineering from Mumbai University, India and a doctorate degree in Computer Science from the State University of New York at Buffalo, USA. His current research is focused on overcoming security and privacy threats in networked computer and cyber-physical systems.

REFERENCES

[1] N. Avlonitis, E. Yeatman, M. Jones, and A. Hadjifotiou. 2006. Multilevel amplitude shift keying in sispersion uncompensated optical systems. *OptoElectronics.*

[2] G. E. Batista, X. Wang, and E. J. Keogh. 2011. A complexity-invariant distance measure for time series. In *SIAM International Conference on Data Mining.*

[3] L.J. Latecki, Q. Wang, S. Koknar-Tezel, and V. Megalooikonomou. Optimal subsequence bijection. 2007. In *IEEE International Conference on Data Mining* (ICDM).

[4] A. Maiti and M. Jadliwala. 2019. Light Ears: Information leakage via Smart Lights. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies.*

[5] M. Muller. *Information Retrieval for Music and Motion*, Springer, 2007.

[6] R. Puri and K. Ramchandran. Multiple description source coding using forward error correction codes. 1999. In *Conference Record of the Thirty-Third Asilomar Conference on Signals, Systems, and Computers.*

[7] M. Wollmer, M. Al-Hames, F. Eyben, B. Schuller, and G. Rigoll. A multidimensional dynamic time warping algorithm for efficient multimodal fusion of asynchronous data streams. 2009. *Neurocomputing.*