Exfiltrating Personal Data from IoT devices

Brandie Pawlowski, Edris Rahimi, Salih Musap Işik, Dailion Janga, Gabriel Lepinay, Yasen Alchev
Cyber Security Advanced
Fontys University of Applied Sciences
Eindhoven, Netherlands
Coach: Mark S. Madsen

***Abstract*–Internet of Things (IoT) devices have privacy risks as they collect, transmit, and store sensitive personal information. Although manufacturers assure privacy protection, the actual behavior of these devices do not always align with their privacy policies. This study examines a range of IoT devices to identify data privacy concerns that apply to these devices, assesses the accuracy of their manufacturer's privacy statements, and determines the technical challenges in evaluating privacy practices of IoT devices.**

Keywords–IoT, privacy, data exfiltration, network analysis, privacy policies, cybersecurity.

I. INTRODUCTION

A. Background on IoT and Privacy Concerns

The Internet of Things (IoT) encompasses interconnected digital devices that exchange data over networks, often without explicit user consent or awareness, examples include smart home assistants, medical devices, e-readers, wearables, and surveillance cameras. Although IoT devices offer convenience and enhanced efficiency, they simultaneously pose significant privacy risks by capturing sensitive data such as voice recordings, health information, location details, and user behavior patterns.

This extensive, often unnoticed data collection raises concerns regarding data security, user consent, transparency, and compliance with privacy laws and regulations. Most users will not have a strong understanding about the nature of data collected, processed, or potentially misused. This emphasizes the need to evaluate the actual practices of IoT devices against manufacturers stated privacy policies.

1) *Relevance and urgency:* Addressing IoT privacy concerns is important due to the growth and widespread usage of IoT devices.

Recent security breaches and unauthorized data disclosures have exposed vulnerabilities, which highlights critical issues around personal privacy and compliance with laws and regulations. For example, one study by computer scientists at the University of California provided research into cross-device user tracking (for when a user might switch from their television to their mobile phone) from TV manufacturers and advertisers. Automatic content recognition (ACR) is used for taking a screenshot of the television display every 10 milliseconds, which is then compared with a database of all content available on Smart TV to determine the user's activity, and then this data is used to cater advertisements towards the user [1].

In a recent example, the Federal Bureau of Investigation (FBI) issued a warning that IoT devices could be compromised through a botnet, BADBOX 2.0. This was found to be a backdoor installed on Android devices before they ever reached the customer [2]. The implications for a home network with an infected device could involve a gateway for attackers to access other connected devices on the network, devices being used as part of a botnet to conduct cyberattacks without the owner realizing.

Inconsistencies in manufacturers' privacy policies result in lost trust in these IoT products. This investigation of IoT privacy practices intends to bring attention to these issues and highlights the importance of safeguarding user data and overall responsibility from the manufacturers in meeting the requirements set by laws and regulations, as well as their own privacy-related policies.

B. Problem Statement

The growing number of IoT devices in homes and daily life has raised serious concerns about data privacy and user consent. Many IoT devices transmit personal information over the internet, often without strong encryption or clear permission from the user. Even though manufacturers claim certain data practices in their privacy policies, studies have shown that these devices may share more data than they admit. This creates risks such as data leaks, unwanted surveillance, and misuse of personal information. The main issue is the difference between what manufacturers say in their privacy policies and what their devices do [3].

C. Research Goals and Scope

TABLE I. MoSCoW TABLE

Goal	Priority
Analyze and document the exfiltration of personal data from at least two IoT devices.	Must
Analyze the collected data to determine whether it aligns with the device's stated privacy policies.	Must

Identify and document any discrepancies or security weaknesses	Should
Perform network traffic and reverse engineering analysis to uncover hidden data flows.	Should
Suggest mitigation strategies or best practices for improving device security.	Could

1) Scope of research: the project includes:

- a review of device privacy policies to understand what manufacturers claim about data collection;
- setup and testing of at least two IoT devices;
- traffic analysis to monitor and capture data transmissions from these devices;
- security assessment of data storage and transmission practices;
- comparison of findings with privacy policy claims to identify potential violations or concerns;
- documentation of results and ethical considerations in a final research report.

II. RESEARCH QUESTIONS

A. Main Research Question

To what extent do Internet of Things (IoT) devices exfiltrate personal data and do the manufacturers abide by their own personal policies?

B. Sub-question

Q1: What do IoT device manufacturers say about data collection and sharing in their privacy policies, and are these statements clear or misleading?

1) Q1 methodology: The methodologies applied for this research question included:

TABLE II. Q1 TRIANGULATION METHODS

Triangulation Method	Description
Library – Literature study	A review on how much personal data is being leaked from IoT devices was performed to establish a baseline understanding of common IoT privacy claims.
Field – Document analysis	A direct analysis was performed on the privacy policy documents of the different IoT device manufacturers (e.g., Fitbit, Amazon Kindle).
Showroom – Ethical check	An ethical evaluation was performed to verify if the manufacturers stated policy aligns with ethical standards and the findings.

Q2: What personal data do IoT devices send out, and how can we detect and analyze this?

2) Q2 methodology: The methodologies applied for this research question included:

TABLE III. Q2 TRIANGULATION METHODS

Triangulation Method	Description
Library – Best good and bad practices	Existing guidelines and documented best practices for ethical IoT data collection and analysis were reviewed.
Lab – Security testing	Device analyses were executed within a controlled setting, using network monitoring tools such as Wireshark and Tshark.
Field – Exploratory data analysis	Captured network traffic from IoT devices was analyzed using manual exploration techniques. For example, the frequency and destination of DNS requests were mapped over time.

Q3: What technical barriers make it difficult to study IoT data leaks, and how can we get around them?

3) Q3 methodology: The methodologies applied for this research question included:

TABLE IV. Q3 TRIANGULATION METHODS

Triangulation Method	Description
Library – SWOT analysis	An analysis on strengths, weaknesses, opportunities, and threats was performed based on the literature study on “How Much Personal Data is Leaked from IoT Devices?".
Lab – Component test	Individual components from IoT devices were tested in a controlled environment. The remarkable 2 was physically connected and its specific background processes were identified and tested.
Workshop – Root cause analysis	During the sprint retrospectives, issues such as failed interception of encrypted traffic (e.g., kindle blocking mitmproxy) were discussed in detail.

III. DEVICE SELECTION

A. IoT devices selection

1) ***The UseeEar***: is an endoscopic camera that handles sensitive health imagery.

As a low-cost IoT device from platforms like Temu, it represents consumer-grade medical equipment that may compromise security while processing intimate medical data.

2) ***Google WiFi***: is selected to represent a major tech company's approach to IoT security.

3) ***IP camera Foscam F19816P***: was selected as a security monitoring device, to see if data related to building security are leaked in any way. Especially using a Chinese product.

4) ***The Kindle Paperwhite***: was selected to represent an e-readers and an IoT device from a major technology company, namely Amazon.

5) ***reMarkable 2***: is a digital paper tablet designed for reading and writing tasks.

Through its proprietary "Connect" cloud service, content is automatically synced across companion mobile and desktop apps, providing real-time access across devices [4]. The connected services include, cloud syncing, WiFI access, and updates; all qualify the remarkable 2 as an IoT device.

6) ***Fitbit Versa Lite***: is a budget smartwatch that offers core smartwatch and fitness features, such as heart rate monitoring, step counting, sleep tracking, and smartphone notifications.

7) ***Google Home Mini:*** an affordable voice assistant device known for integrating deeply with Google's data-driven services.

IoT devices like smart speakers offer convenience but also introduce privacy risks. These devices often send data to cloud services, with the extent of this data transfer and its alignment with privacy policies remaining unclear to end users. This device was selected to represent the low budget version within the smart speakers.

8) ***Apple HomePod Mini***: Apple's voice assistant device focused on privacy-centric architecture and local data processing.

This device was selected to showcase a higher and version within the smart speaker group to see whether it differs from the budget versions when it comes to privacy claims etc.

9) **Google Nest Hub**: a smart display IoT device that combines the functionality of a smart speaker with a touchscreen display, allowing it to show visual information alongside voice responses.

B. Manufacturer Claims from Privacy Policies

1) ***The UseeEar***: manufacturer [5] claims "reasonable security measures" and states data won't be shared with third parties for marketing without consent.

2) ***Google WiFi***: is clearer and more specific compared to many other IoT devices. It mentions that no browsing history is being tracked, only basic WiFi performance data like signal strength and device types.

While personal info is being collected during initial setup, google claims that its not linked to the user. Additionally, user data is stored for 180 days, and account info stays until the user deletes it. Lastly, they claim that personal data is not being sold to third party vendors.

3) ***IP camera Foscam F19816P***: Shenzhen Foscam Intelligent Technology Co., claims that all data are processed and stored in China [5]. They claimed that the information collected is the one they provide during registration: Name, email, username, camera info, IP. And they automatically collect IP addresses, cookies, browser data, and pages visited. Regarding data sharing, Foscam does not sell or rent data, but may share it with advertisers, affiliated businesses and business partners.

4) ***The Kindle Paperwhite***: Amazon's privacy disclosures for Kindle devices are vague and use general terms.

Amazon collects device and usage data, such as reading habits and time spent on each page but does not explain in detail how this data is processed [6].

Using the Kindle requires linking it to an Amazon account, tying personal identifiers to all activity. Data retention periods are not specified, and data appears to be stored indefinitely unless the user takes action to delete it [6]. While Amazon states it does not sell personal data, it shares information with affiliates and service providers [6].

5) **reMarkable 2**: claims that all data from the device, mobile, desktop apps, or browser is encrypted in transit (TLS) and at rest (AES-256) [4] [7]. The manufacturer states compliance with data protection frameworks, including GDPR and CCPA. For device security, the tablet supports a six-digit passcode, optional two-factor authentication for accounts, secure boot features, and on-device encryption.

6) **Fitbit Versa Lite:** Fitbit is part of google, meaning that most of google privacy policy applies to the smart watch as well.

Fitbit claims transparent data collection and why it is used [8]. There are some optional features like menstrual tracking or glucose logging. Overall, it comes down to, anything that is needed for normal watch usage, is used. The manufacturer explicitly claims they will not sell or tie personal data of its users.

7) ***Google Home Mini:*** according to Google's privacy documentation, the Home Mini is designed to only listen for the "Hey Google" wake word and claims not to record or send anything until it hears that. They also state that most of the processing is done on the device itself, and only necessary data is sent to Google's servers to improve responses or maintain functionality. In terms of stored data, Google allows users to manage their voice activity via the Google account dashboard and says the user has full control over what is kept or deleted.

8) ***Apple HomePod Mini***: Apple's privacy policy focuses a lot on minimizing data collection and doing as much processing on-device as possible. For example, Apple says that Siri recordings are not stored by default unless the user explicitly allows it. They also claim that queries are anonymized before being sent to Apple servers and not linked to a personal Apple ID.

HomePod Mini uses something Apple calls a "random identifier" instead of your actual user info when sending requests to their servers. In terms of security, Apple highlights that all traffic is encrypted and the device supports features like secure boot and hardware-based encryption. According to them, even Apple can't access certain data because it's encrypted end-to-end. Compared to Google, their statements are more focused on local control and not monetizing user data.

9) **Google Nest Hub:** Users can review and delete stored voice recordings and old data via the Google Home app or web interface.

The device listens for the wake word ("Hey Google") and only sends audio recordings to Google's servers after the wake word is detected.

Google claims to encrypt data in transit and at rest, and states that user data is not sold to third parties.

Users are provided with multiple privacy settings, including controls over voice activity storage and the ability to mute the device's microphone.

C. Clarity and Consistency of Claims

1) *The UseeEar*: policies contain many contradictions. For example, it mentions the use of "reasonable security measures" to protect data, but it doesn't state what these measures are.

Additionally, it is unclear how data is being processed, since the policy only mentions "third-party providers" without providing specific names.

2) *Google WiFi*: is more transparent than most IoT manufacturers when it comes to privacy and security.

The use of WPA3, automatic software updates, and even the tracking of known vulnerabilities, shows they care. However, despite these strong privacy claims, traffic analysis still shows many DNS queries to advertising networks, which raises questions about how private the device really is.

3) *IP camera Foscam F19816P*: There is an unambiguous state that under 18 users are not allowed and that such data will be deleted if discovered.

The policy says users have control over their data but also says: "some information... in our private records after deletion." So, this is inconsistent with the claim of respecting user deletion requests.

The policy says, "If your use... violates any law applicable to you... your right to use the Services is revoked." This shows that all legal risks are to the user, without providing legal & safe features for international users.

They say the claims to protect data but then adds: "We cannot guarantee complete security." Honestly, this is vague and lacks clear details on what protections are in place. They say they cannot control how other users share your uploaded content.

4) *The Kindle Paperwhite*: Amazon provided limited transparency.

There is no public vulnerability disclosure policy, detail on firmware update cycles, no accessible CVE tracking or security bulletins specific to Kindle. User control over privacy is minimal, mostly limited to toggling ad personalization and basic data sharing options; more advanced controls, like disabling telemetry are not available [9].

5) reMarkable 2: privacy policy emphasizing user control and transparency, forensic analysis of the device reveals something else.

Background services such as crashuploader and memfaultd automatically collect and transmit telemetry data, including crash logs, and the device serial number

without user consent. These services operate quietly in the background and are enabled by default, with configurations set to upload data every 20 minutes.

6) ***Fitbit Versa Lite***: Fitbit doesn't sell personal data, that does not mean that developers cannot access it. Data can be retrieved through API if a user has a third-party app installed. Optional features are indeed not passed through API if they are not enabled.

Fitbit still has a duty to make sure that user data is not stolen through this exploit according to article 5 (data protection principles) from the GDPR [10].

7) **Google Home Mini**: while reading Google's privacy documents, some things seem clear on the surface, but once you look closer, it gets kind of confusing. For example, they say the Home Mini only sends data when it hears "Hey Google," but in our tests the device was clearly still talking to Google's servers even when it was just sitting there doing nothing. There were regular DNS requests and other encrypted traffic, which makes it hard to trust the "only when necessary" claim.

Another thing is that they say users can manage their data in their Google Account, but it's not always obvious what kind of data is being collected in the first place. There's no full list of what gets sent or how often. Also, they state that they don't sell personal data, but they still use it for ad personalization, which feels a bit like a loophole. So, in general, the privacy policy sounds good, but the real behavior doesn't always match.

8) **Apple HomePod Mini**: Apple claims seem more consistent with what we saw. The HomePod Mini didn't show any sketchy connections or ad-related domains in the traffic. It mostly just contacted Apple's own infrastructure, and not too often either. That kind of lines up with their promise to keep most data processing on the device and avoid unnecessary tracking.

But even with Apple, there's stuff that's not 100% clear. For example, they talk about anonymization and random identifiers, but they don't really explain how that works technically. Also, it's not easy to find out what exact data is collected when you use Siri or iCloud sync on the HomePod. There's also no public-facing telemetry dashboard or something like that, so users just must trust that Apple handles everything the way they say they do.

9) **Google Nest Hub**: Google's privacy statements are relatively clear. However, there are still a few things that aren't clear.

The full scope of data collected beyond voice interactions is not described in detail, including sensor data, screen usage, and other data logs.

It is not always clear how long all types of data are kept, and when data might be shared for analytics or to improve the product.

There's no in-depth explanation of how data is minimized, anonymized, or handled by third parties.

IV. FINDINGS & ANALYSIS

A. The UseeEar

The UseeEar (Fig. *1*) device presented critical security vulnerabilities, creating an open, unencrypted access point for medical examination streaming. Successful ARP poisoning attacks via Ettercap enabled complete traffic interception and reconstruction of MJPEG video streams containing sensitive medical imagery. The device communicates with suspicious Chinese infrastructure (yun.simicloud.com), though packet loss prevented analysis of external data transmission content. This consumer medical device violates fundamental privacy principles by transmitting intimate medical examinations without encryption, creating serious regulatory compliance risks and demonstrating vulnerability to man-in-the-middle attacks.

Technical challenges and workarounds: VMware network bridging issues caused packet loss during external server communication analysis. Video stream reconstruction required format identification (MJPEG vs H.264) and custom extraction techniques using tshark and ffmpeg to successfully decode intercepted medical footage.

Fig. 1. Endoscopic Camera UseeEar

B. Google WiFi

Network traffic analysis revealed the Google Home device (Fig. *2*) implements strong encryption practices, with nearly all communications secured via TLS/SSL. The only unencrypted traffic consisted of certificate validation requests (OCSP), which are legitimate security features. However, significant privacy concerns emerged through extensive DNS queries to advertising networks (googleads.g.doubleclick.net, ssl.google-analytics.com) and third-party tracking services, indicating comprehensive user behavior monitoring. While no malicious data exfiltration was detected and the device follows modern IoT security standards, the volume of advertising-related communications raises questions about data collection practices.

Technical challenges and workarounds: Standard encrypted traffic required DNS analysis and connection pattern examination rather than payload inspection to understand device behavior.

Fig. 2. Google WiFi

C. *IP camera Foscam F19816P*

The Foscam IP camera (Fig. *3*) device didn't present vulnerabilities. The product was developed following the standards. Even if the product is discontinued, it supports WPA2 encryption, remote access via a mobile app, protected with accounts and credentials. The only info extracted where the DNS request to Foscam services, and the mac address, camera name and device ID that are broadcast on the LAN.

Fig. 3. Foscam IP camera F19816P

D. *The Kindle Paperwhite*

The Kindle Paperwhite (*Fig. 4*) was not found to have any critical vulnerabilities throughout the privacy analysis. The traffic observed was consistent with expected behavior from an Amazon device, with several DNS queries to various amazon services. The traffic was encrypted. Network scans did not reveal any open ports, and the Kindle did not response to nmap probes. Passive observation showed that the Kindle announces itself on the local network using ARP, which is common for any connected device (*Fig. 5*).

Direct traffic capture through Wireshark revealed ARP and multicast activity, but no sensitive data was transmitted in plaintext. Attempts to intercept or proxy network communication using tools like mitmproxy and Fiddler were unsuccessful, due to the Kindle blocking internet access when certificate interception is detected. This finding would suggest that Amazon uses certificate pinning and secure DNS resolution. No user data or session activity could be monitored without jailbreaking the device.

DNS queries and traffic patterns indicate routine connectivity to Amazon services, including content delivery (a4k.amazon.com), sync and telemetry (unagi-na.amazon.com, cde-ta-g7g.amazon.com), and cloud storage (Amazon S3 and CloudFront domains) (*Fig. 6*). While this does confirm that the device regularly reports to Amazon infrastructure, all data was transmitted over encrypted HTTPS, with no indication of unnecessary data leakage on the local network.

However, the extensive list of Amazon domains contacted by the Kindle during normal use raises privacy concerns for the average user. While all communication appears encrypted, the sheer volume and variety of endpoints, which range from telemetry (unagi-na.amazon.com, cde-ta-g7g.amazon.com), to device messaging and content delivery, suggest that the device is continuously syncing data and reporting back to Amazon. For a privacy-conscious user, it could be unsettling to know that the Kindle interacts with numerous services beyond the core reading function, potentially sharing usage patterns, device status, behavioral metrics, and reading habits. Without transparent controls over what data is collected or how long it's retained, users are giving up more personal information than necessary from using a device marketed for offline reading.

Fig. 4. Kindle Paperwhite 16GB

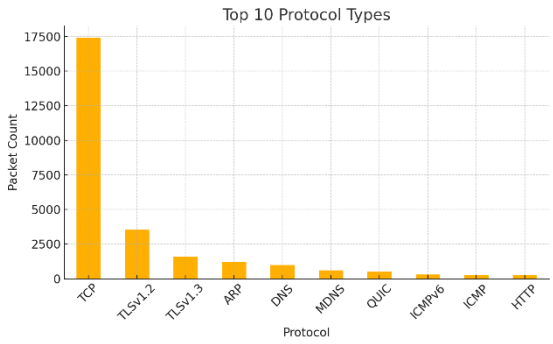

Fig. 5. Top protocol types - Kindle Paperwhite

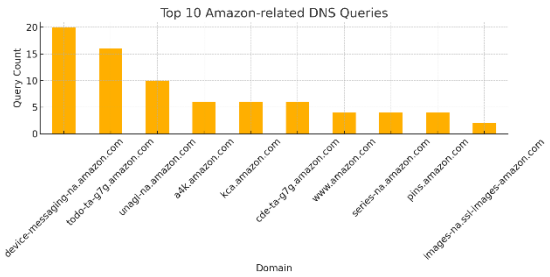

Fig. 6. Amazon DNS queries - Kindle Paperwhite

E. reMarkable 2

The reMarkable was connected to a pc using a type C cable, it received a local IP address (10.11.99.1) and allows SSH access as root using a randomly generated password created during initial setup. This granted full access to the device's operating system, which is a minimal BusyBox Linux Environment.

By using the ps command, several notable background services were identified:

1. /usr/bin/crashuploader, when executed, the binary collects the device's serial number and attempts to upload crash data (including the serial and location) to a Remarkable owned endpoint.
2. /usr/bin/memfaultd, is a telemetry collection client that collects for example, device hearbeat (*Fig. 7*). The system checks if the user is using the tablet every 20 minutes and performs online checks (user connected to the internet) as well.
3. /usr/sbin/collected: Likely another telemetry logger, this process references a config file at /etc/collected.conf, though its exact function remains unclear.

Manual testing confirmed that telemetry uploads can be disabled by editing the configuration file for memfaultd and setting "enable_data_collection": false. After modifying this setting, stopping the service, it prevents it from restarting on reboot.

The findings confirm the presence of telemetry activity. The device stores and uploads various types of diagnostic data on a recurring basis.

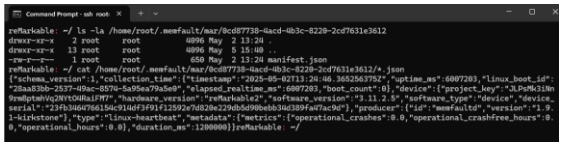

Fig. 7. JSON output for a Linux-heartbeat event

F. Fitbit Versa Lite

The Fitbit Versa Lite (Fig. *8*) does not have a direct WiFi connection and relies on the mobile phone it is connected to. Through Fitbit's developer portal, it was possible to register a fake application and gain access to sensitive user data, such as profile details, sleep patterns, and heart rate using only access tokens. The access tokens are generated and given to the developer when a user installs the fake application.

This demonstrates that legitimate looking third party apps on Fitbit can access personal data while google claims not to sell user data.

The Privacy breach is not in the device itself, but in the app ecosystem that supports it. The combination of broad API permissions and weak vetting of app developers makes it easy to misuse sensitive health data under the guise of legitimate functionality.

Fig. 8. Fitbit Versa Lite

G. Google Home Mini

Network traffic analysis revealed that the Google Home Mini (*Fig. 9*) implements strong encryption practices, with nearly all communications secured via TLS/SSL. The only unencrypted traffic consisted of certificate validation requests (OCSP), which are part of the standard security protocol. Despite being idle, the device frequently initiated outbound DNS queries to domains such as googleads.g.doubleclick.net, ssl.google-analytics.com, and other advertising/tracking services (*Fig. 10*). This suggests that even during periods of inactivity, the device performs background tasks that involve user behavior profiling and telemetry synchronization [11] [12].

The payload sizes of outbound communications varied, with some packets exceeding 1.2KB. These bursts of communication occurred intermittently and were not user-initiated, contradicting Google's policy statements that claim the device only communicates when necessary. The DNS metadata indicated regular contact with both core Google infrastructure and third-party analytics endpoints. This discrepancy highlights a lack of

transparency in actual device behavior compared to what is described in privacy documents [11].

Technical challenges included the inability to inspect packet contents due to TLS encryption. To compensate, we relied on DNS traffic logs, domain resolution patterns, and analysis of communication intervals. Additionally, because of the closed nature of the firmware, no direct debugging or packet injection was feasible without rooting the device.

Fig. 9. Google home mini

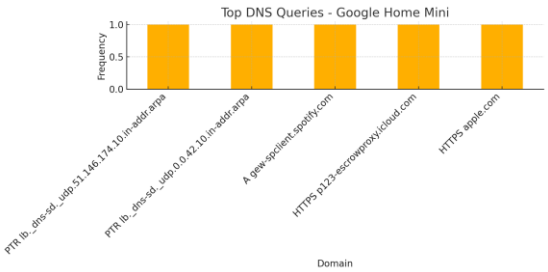

Fig. 10. Top DNS queries - Google Home Mini

H. Apple HomePod Mini

The Apple HomePod Mini (*Fig. 11*) demonstrated a significantly quieter communication profile. It maintained strong TLS encryption and only communicated with Apple domains such as *.icloud.com and *.apple.com. DNS traffic was sparse compared to the Google Home Mini, and we observed no connections to third-party advertising or analytics services. This aligns with Apple's public commitment to privacy-focused design [13] [14].

Most observed payloads were small (under 512 bytes) and occurred periodically, indicating background syncing or system-level status checks. During testing, no user-initiated queries were captured outside of setup, further supporting Apple's claim of minimal and necessary data collection (*Fig. 12*). Certificate pinning and system-level encryption blocked traffic interception tools like mitmproxy, forcing reliance on passive observation.

Technical challenges included Apple's use of aggressive encryption and device sandboxing, which prevented packet decryption. Attempts to analyze live session traffic were met with connection rejections due to invalid certificates. The analysis was therefore limited to destination IPs, frequency of requests, and reverse DNS lookups of contacted servers.

Fig. 11. Apple HomePod Mini

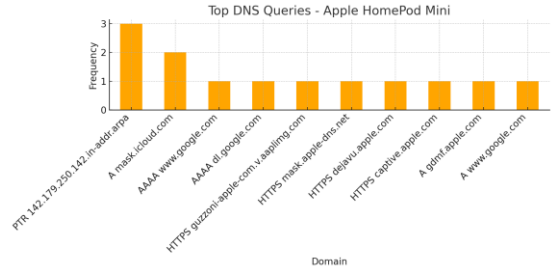

Fig. 12. Top DNS queries - Apple HomePod Mini

I. Google Nest Hub

All traffic from the Nest Hub (*Fig. 13*) was encrypted with TLS, matching Google's claims regarding secure data transmission. The DNS queries showed that the Nest Hub communicates regularly with various IP's. Including Google's services, such as DNS servers (8.8.8.8 and 8.8.4.4) and other services like advertising-related domains (doubleclick.net). This ongoing background communication happened even when the device seemed to be idle. A significant amount of data was sent to various IP addresses within a short period, even when the device was idle, which was quite concerning.

No evidence was found of sensitive user data being transmitted in plain text. However, the frequent connections to Google's services indicate that the Nest Hub collects extensive usage metadata, which may include device interactions, schedule routines, and possibly ambient sensor information. Since all the traffic was encrypted, it wasn't possible to see exactly what data was being sent.

The main challenge was the device's use of strong encryption, which made payload inspection difficult without device compromise. Additionally, the lack of physical ports prevented direct hardware-based analysis. As a result, the research focused on network metadata,

such as source and destination address, traffic frequency, and timing patterns.

Fig. 13. Google Nest Hub

V. CONCLUSION

This research aim was to uncover how Internet of Things (IoT) devices handle personal data and whether their behavior aligns with the privacy claims made by their manufacturers. Through a combination of policy analysis, lab-based network monitoring, and real-time traffic inspection, we discovered that many IoT devices exhibit inconsistent, vague, or misleading privacy practices.

While some devices demonstrate transparent behavior, many others communicate continuously with remote servers, transmit data in ways that were not disclosed in their privacy statements, or collected more information than users likely realize. “Idle” devices were often not truly idle; they sent regular DNS queries, communicated with third-party domains, or uploaded telemetry and crash data without any clear notification to the user.

Several devices stood out for their high level of background activity, such as the Google Home Mini and Google Nest Hub, which sent encrypted but frequent traffic to advertising and analytics services. In contrast, the Apple HomePod Mini showed a more restrained communication pattern, limited to Apple’s own domains. However, even devices marketed as offline or privacy-focused, like the Kindle Paperwhite or reMarkable tablet, transmitted regular telemetry data, raising concerns about transparency and user awareness.

In the most extreme case, the UseeEar endoscopic camera transmitted sensitive medical data over unencrypted connections, creating significant regulatory and ethical risks. Meanwhile, the Fitbit Versa Lite highlighted how easily personal health data can be accessed through third-party apps, depending on user permissions that are often granted without understanding the consequences.

The findings have one big common pattern. There is a growing disconnect between what manufacturers say, and what their devices do. Even when the privacy policies are (technically) accurate, they often rely on ambiguous language or omit important details that would help users make informed decisions. Encryption is widely used, but it is not a substitute for responsible data handling or meaningful transparency.

This paper also revealed some technical challenges. TLS encryption, certificate pinning, and closed firmware environments limited our ability to analyze payloads. As a result, much of our analysis depended on DNS traffic, metadata patterns, and passive observation. These methods proved effective in identifying data exfiltration trends but also highlight the difficulty of holding manufacturers accountable without access to the full data path.

In conclusion, as IoT adoption continues to grow in popularity and capability, their privacy behaviors often fall short of what users reasonably expect. To address this, manufacturers should secure data with encryption, limit unnecessary data collection, be clear and honest in their privacy policies, and provide users with real control over their personal information. Independent research like this remains essential in verifying claims and encouraging the industry to adopt more transparent and privacy-conscious standards.

REFERENCES

[1] A. Fell, "Your Smart TV is Watching What You Watch," UCDavis - College of Engineering, 18 December 2024. [Online]. Available: https://engineering.ucdavis.edu/news/your-smart-tv-watching-what-you-watch. [Accessed 12 June 2025].

[2] K. Beek, "BADBOX 2.0 Targets Home Networks in Botnet Campaign, FBI Warns," DarkReading, 6 June 2025. [Online]. Available: https://www.darkreading.com/threat-intelligence/badbox-home-networks-botnet-campaign-fbi. [Accessed 12 June 2025].

[3] P. Menard, "Analyzing IOT users’ mobile device privacy concerns: Extracting privacy permissions using a disclosure experiment," ScienceDirect, 3 January 2020. [Online]. Available: https://doi.org/10.1016/j.cose.2020.101856.

[4] reMarkable, "reMarkable 2," [Online]. Available: https://remarkable.com/products/remarkable-2. [Accessed 01 04 2025].

[5] Shenzhen Yipincheng Technology Co.,LTD, Shenzhen Yipincheng Technology Co.,LTD., 2014. [Online]. Available: https://yipincheng.en.ec21.com/company_info.html.

[6] Amazon, "Amazon.com Privacy Policy," Amazon, 14 February 2025. [Online]. Available:

https://www.amazon.com/gp/help/customer/display.html?nodeId=GX7NJQ4ZB8MHFRNJ. [Accessed 10 June 2025].

[7] reMarkable, "Data security," [Online]. Available: https://support.remarkable.com/s/article/Data-security.

[8] Fitbit LLC, "Fitbit Privacy Policy," Google, 16 September 2024. [Online]. Available: https://support.google.com/product-documentation/answer/14815921?hl=en.

[9] Amazon, "Kindle E-Reader Software Updates," Amazon, 2025. [Online]. Available: https://www.amazon.com/gp/help/customer/display.html?ref_=hp_left_v4_sib&nodeId=GKMQC26VQQMM8XSW. [Accessed 10 June 2025].

[10] E. U. Law, "GDPR," EUR-Lex, [Online]. Available: https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng.

[11] Google Support, "Google Nest Privacy Help," Google, [Online]. Available: https://support.google.com/googlenest/answer/7072285. [Accessed 20 May 2025].

[12] Google Account Help, "Your data in the Assistant," Google, [Online]. Available: https://myactivity.google.com/product/assistant. [Accessed 20 May 2025].

[13] Apple, "Privacy Overview," Apple, [Online]. Available: https://www.apple.com/privacy/. [Accessed 11 March 2025].

[14] Apple, "Siri Privacy and Data Handling," Apple, [Online]. Available: https://support.apple.com/en-us/HT210657. [Accessed 11 March 2025].