Applied GIS

ISSN: 1832-5505

Vol-11 Issue-01 JAN 2023

The Effect of Internet-of-Things Locations on MUDs

RAMRAJSINGH PRATHAP SINGH1, CH SWAPNA

Assistant professor1,2

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

P.B.R. VISVODAYA INSTITUTE OF TECHNOLOGY & SCIENCE S.P.S. R NELLORE DIST, A.P, INDIA, KAVALI-524201

Abstract—

Understanding the domains, protocols, and ports that an IoT device communicates across is a basic barrier for IoT security and identification. Analytical and management solutions in these domains need to be able to identify and authenticate devices, as well as understand what constitutes regular device activity. Manufacturer Usage Description (MUD) is a white-list protection technique developed by the IETF; it uses a MUD file to codify permitted network activity, which can subsequently be used as a firewall. We show that it is harder than one may think to understand what is typical behaviour for an IoT device. The location of an Internet of Things (IoT) device may affect how it operates, the domains it can connect to, and even the protocols it uses. We break down and explain scenarios in which device behaviour is affected by geography. Next, we demonstrate how MUD files may be simplified in a more general sense. We can construct a universal MUD file that can be used in any region by processing MUD files from a variety of regions. We built MUDIS, a MUD file comparison and generalisation tool, to undertake the study. In order to help academics and IoT manufacturer's view, analyse, and generalise MUD files, we have made our MUDIS tool and dataset publicly accessible online.

INTRODUCTION

The Internet of Things (IoT) industry is fragmented, with several players and no overarching standard for how devices should be built, protected, and recognised on the network. Specifically, this study looks at how the physical location of a device affects its behaviour in a network. We discovered that the exact same Internet of Things (IoT) device, although having identical firmware, exhibited varying degrees of behaviour and communicated with a wide variety of domains, protocols, and ports depending on its physical location. As far as we're aware, this is the first piece of study to formally identify device location as a factor influencing device behaviour. For this analysis, we combined the data from Ren et al. [1], who recorded devices that were both physically placed and logically connected in two locations, with our own dataset containing measurements for devices in our lab that were virtually connected to different locations using VPN or logically connected to different locations by registering the device in the IoT application in different countries. We demonstrate that the physical location of an Internet of Things (IoT) device will affect its network behaviour for a variety of reasons. One example is that for marketing purposes, the same IoT may have various characteristics depending on its location. Other examples include insufficient encryption, privacy rules, CDN-like solutions, and country-specific legislation. Existing literature on this topic is limited to a discussion of how data protection laws in the United Kingdom and the United States (GDPR, FTC) affect the online behaviour of Internet of Things devices [1]. Our research, on the other hand, examines the role of location across a

wide range of nations and shows that other factors also contribute to the observed disparities. The Manufacturer Usage Description and the security of the Internet of Things are both directly affected by this occurrence (MUD). As an IETF standard [2], the MUD allows us to codify the proper conduct of IoT gadgets on the network. The MUD file checks whether or not the device is being hacked in a fashion that is similar to an Access Control List (ACL) or a network firewall. Since the IoT device may use DHCP or LLDP to get the MUD file, just a single MUD file is needed for each firmware version, independent of the device's physical location. However, our research reveals that, in 90% of tested devices, the same device with the same software exhibits distinct network behaviour in various geographic regions. We observed that in many circumstances, the device's network behaviour is not determined by its actual physical location but rather by the logical geographic region selected by the user during account registration (i.e., using geo IP)

BACKGROUND MUD

The MUD serves a dual purpose in our methodology. Initially, we can study network activity at the flow level since the MUD file formalises that information. Second, MUD technique is a security solution, and enhancing it is

ISSN: 1832-5505

Vol-11 Issue-01 JAN 2023

a primary goal of this project. The MUD Internet Standard [2] is designed to describe the proper traffic patterns for IoT devices, which in turn reduces their attack surface. Any data that doesn't fit this profile is likely harmful and should be prevented in some way. Manufacturers of IoT devices give these details in MUD files. All MUD files are made up of Access Control Lists (ACLs) that have numerous Access Control Entries (ACEs). Figure 3 depicts how each ACE is defined as a 5-tuple.

$$ACE = (legitimate\_endpoints, protocol, source\_port, destination\_port, direction) \quad (1)$$

The IoT's valid endpoints are the locations to which it may send and receive data. Common methods for defining these include the use of domain names or ranges of domain names [2, 18] (like *.iotvendor.com), IP subnets (including *), and media access controls (MACs) for local area networks. In this regard, we point out that the MUD [2] specification strongly suggests substituting domain names for IP addresses.

Accept or drop are the common ACE responses. Because of the whitelist in the MUD file, any communication that does not match an ACE is automatically dropped. To lessen the vulnerability of a firewall or AAA server, for example, the MUD manager reads the MUD file and deploys the appropriate Access Control List (ACL) rules. As a result, manufacturers have the difficult challenge of creating a full and representative MUD that accounts for a wide variety of factors, including the usage of third-party libraries, the OS network behaviour, the whole device's operational activities, and more. Tools that create MUD files from network captures [3, 10] are available to help with this problem. Another strategy involves a network security component [19] that collects the MUD file from real-world traffic and analyses it using big data. When dealing with IoT suppliers that lack the resources or motivation to produce a MUD file, this is a useful workaround. Analysing the Effects of Device LocationCaptured network traffic data (cap files) from our lab's router and log files from Ren et al. [1] make up our dataset. The 31 IoT devices (plugs, cameras, light bulbs, and so on) used in our captures are situated in up to 14 different countries (remotely or virtually) and make full use of all of their features [20]. Activation locations were determined by whether nations supported the IoT user application's registration and provisioning procedure. We discovered that the nation selected during device provisioning had a far greater impact on the device's network behaviour than the device's actual physical location (as shown by the IP address of the device as viewed in the VPN) (see full technical report for more details [21]). Moving on, we used MUDGEE [3] to transform the pcaps into MUD files. You may get the whole list of tested devices and the associated MUD files organised by country at [8]. 2

The number of shared ACEs between two MUDs is divided by the sum of all shared ACEs to get the Jaccard similarity coefficient, which may be used to compare and contrast the two MUDs. Let's call the MUD of device d at position I MUDd i. Two MUDs for the same device (device d) at different locations I j) are comparable if and only if:

$$Similarity_d(MUD_i^d, MUD_j^d) = \frac{|MUD_i^d \cap MUD_j^d|}{|MUD_i^d \cup MUD_j^d|} \quad (2)$$

Figure 1 displays the CDF of MUD similarity values for the devices in our dataset and compares the generated MUD files across various geographical areas. Nearly 80% of MUD comparisons exhibit similarity measures lower than 0.7, indicating that device location has a major influence on the MUD.

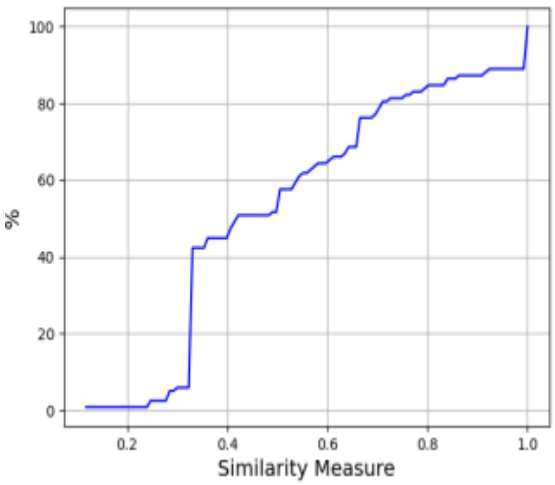


Figure 1: Cumulative Distribution Function (CDF) of MUD files similarity scores for all the devices in the dataset. Each similarity score is calculated by comparing two different locations MUD files of a device. Each device was captured in up to 14 locations.

Applied GIS

ISSN: 1832-5505

Vol-11 Issue-01 JAN 2023

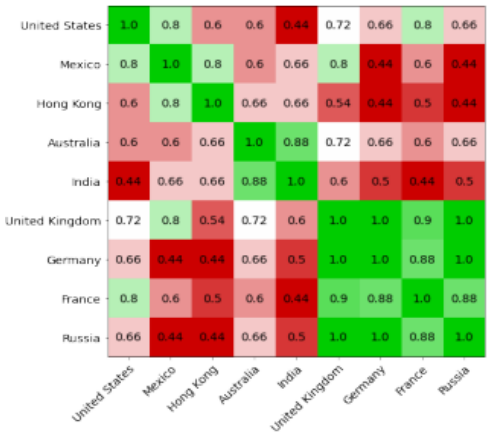


Figure 2: Heat map of similarity measure for the Yi camera, across ten different logical locations. The heat-map highlights that cross-region locations have lower similarity scores.

In Figure 2, we take a deep dive and focus on an individual device, investigate its MUD similarity scores as a function of the geographical location. Figure 2 shows the MUD similarity heat-map of the Yi camera MUD files as measured in ten countries. We ordered the countries according to region. As can be observed, locations further away from each other (cross-regions) have lower MUD similarity values. Throughout our experiments, we observed that some device functionalities were not supported in all locations. For example, the Xiaomi camera face recognition features were supported only in the Chinese region. The reasons range from

Table I shows a comparison of the network behaviour (power supplies, ports, and protocols) between two logical locations using Xiaomi cameras.

	China	Israel
Domain Names	Fixed IP	sg.ots.io.mi.com
Port	HTTP (80)	HTTPS (443)
IP Resolution	HTTP Request	DNS
Encryption	Self-signature	Standard TLS

from government policies to manufacturer advertising. It is normal practise for a producer to release many varieties of a product, with one for each market (e.g., [23]).

PARALLELS IN THE MUDS

This section compares and contrasts two MUD files. We found that changes to ACEs most often concerned the permitted endpoints' domain names. Our results show that there are subdomain-specific variations for 80% of the devices. To provide just one example, the Samsung SmartThings Hub (shown in Figure 3a) is compatible with two distinct domains in the United Kingdom and the United States: dc-eu01-euwest1.connect.smartthing.com and dc na04-useast2.connect.smartthing.com. However, of the devices in the sample, 9% had a unique top-level domain (TLD). As an example, the Yi camera may exchange data with many top-level domains (TLDs), including api.xiaoyi.com.tw in Hong Kong and api.eu.xiaoyi.com in Germany. Since the user specifies the device's logical location during registration, we infer that the manufacturer may implement various features and rules depending on the domain IDs used.

Keep in mind that the manufacturer can have decisions made based on physical location by employing a standard DNS server that can connect a single domain to different servers, according to the geo-locations; in this case, however, the user would not have the option to select a different logical location. When comparing ACEs from different MUDs, we consider them to be comparable if their domain names3 are similar and all other fields are also similar. We next demonstrate that, by include ranges in the sub-domain name, it is possible to generalise a set of ACEs that are quite similar to a single ACE in the generalised MUD. It's the device's location that matters more in certain situations. In the instance of the Xiaomi camera shown in Table I, for example, the port and protocols utilised by the device vary depending on its location. In our technical paper [21], we explain how we compare ACEs in further depth by separating those that face the same direction of traffic into two groups. It is possible to distinguish between two types of ACEs: (1) those with the same or similar endpoints but a different port or protocol, and (2) those with the same or similar endpoints but a different port and protocol.

FOG CATEGORIZATION

The purpose of this chapter is to build a safe, complete, and reliable MUD that can be used by anybody. The universal MUD should work in all of them, which is what we mean by "comprehensive." Ultimately, the MUD has to be restrictive because

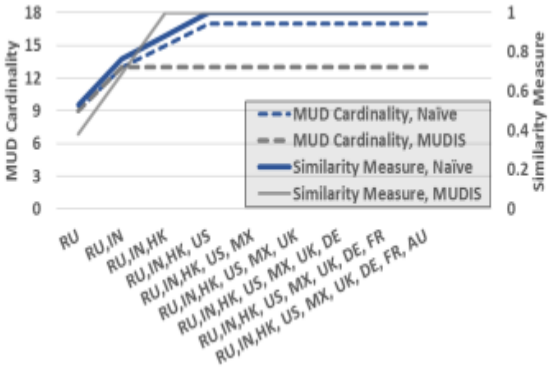
Applied GIS

ISSN: 1832-5505

Vol-11 Issue-01 JAN 2023

Figure 3: Two MUDs of SmartThings hub ACEs, one from the United States (top) and one from the United Kingdom (bottom) (bottom). With the exception of the endpoint being.connect.smartthings.com, MUDIS's generic ACE (right) has the same parameters as the original.

whitelisting allows you to restrict IoT traffic to just authorised sources, making your devices more secure. In order to generalise across two MUDs, the most fundamental generalisation method is multi-MUD. Using an iterative procedure, we may use the technique to generalise n MUD files by feeding the results from the previous iteration into the n the MUD file. With as few iterations as possible, we want to develop a fully-featured and adaptable MUD. In this paper, we demonstrate that our technique achieves faster convergence than the naïve algorithm. Adding additional MUDs from different regions will not alter the overall MUD experience. The scope and density of a naïve generalisation technique that simply unifies all existing MUDs would be optimal. However, we demonstrate that its convergence is sluggish and that it produces a bigger MUD file than our MUDIS generalised MUD. Administrators or manufacturers who must maintain a MUD will like a file with fewer rules since it is easier to explain to humans. Additionally, it provides a lightweight firewall implementation. The concept behind MUDIS generalisation is to combine two ACEs that are otherwise identical save for a single character in the sub-domain field (see Figure 3b for an example,connect.smartthings.com). In order to maintain the integrity of the generic MUD, MUDIS does not gene.



Compare the Yi Camera's generalised MUD and naive unifying MUD file performances in Figure 4. For each place along the x-axis, the unified or generalised MUD is represented by the dot. Each MUD is compared to the global MUD, which includes all possible regions, to get a similarity score.

differentiate ACEs with various TLDs (e.g., iotvendor.co) and well-known, client-shared cloud services (e.g.,s3.amazonaws.com). With respect to subdomains, MUDIS only generalises when the whole domain is under the control of the primary domain owner, i.e., the IoT manufacturer or the precise IoT service that the manufacturer employs. This is consistent with the IETF's DNS for IoT Operational Consideration [24]. If several identical ACEs share a domain that has already been generalised for some of their counterparts, we will generalise that domain for all of them to guarantee rapid convergence. We expand it to handle future locations we haven't met yet based on the crucial insight that such a domain surely contains a segment that relies on the location. The convergence study of MUD files for the Yi Camera from 10 different places is shown in Figure 4. As quick convergence is of primary importance to us, we rank the sites such that we choose those from varying areas first. Locations outside of their respective areas tend to have lower similarity scores, as illustrated in Figure 2. This means that they provide more specific details to the global MUD. We contrasted our generalisation technique for MUDIS to a naïve one that just combines all existing MUDs. Every dot on the x-axis represents a particular generalisation of a MUD in one of the places you've chosen. Point RU, IN, for instance, is the generalised MUD that results from combining the RU (Russia) and IN (India) MUDs. We provide the cardinality (number of ACEs) and similarity score to the associated global MUD for each generalised MUD; the global MUD is the result of running the algorithms (naive or MUDIS) on all of the locations. In terms of both cardinality and

convergence time, our version of the MUD method outperforms the naive approach.

Conclusion

Our results show that the physical placement of a device influences its network activity and the MUD values it generates. To make a universal MUD, we provide a powerful generic programming approach.

REFERENCES

[1] J. Ren, D. J. Dubois, D. Choffnes, A. M. Mandalari, R. Kolcun, and H. Haddadi, "Information exposure from consumer Iot devices: A multidimensional, network-informed measurement approach," in IMC Conference. New York, NY, USA: ACM, 2019, pp. 267–279.

[2] E. Lear, R. Droms, and D. Romascanu, "Manufacturer Usage Description Specification," RFC 8520, Mar. 2019. [Online]. Available: https://rfc-editor.org/rfc/rfc8520.txt

[3] A. Hamza, D. Ranathunga, H. Gharakheili, M. Roughan, and V. Sivaraman, "Clear as mud: Generating, validating and applying iot behavioral profiles," in Workshop on IoT Security and Privacy. USA: Association for Computing, 2018, pp. 8–14.

[4] NIST, "National institute of standards and technology," Sep 2021. [Online]. Available: https://www.nist.gov/

[5] N. I. o. S. NIST and Technology, "Methodology for characterizing network behavior of internet of things devices," Apr 2020. [Online]. Available: https://nvlpubs.nist.gov/nistpubs/CSWP/ NIST.CSWP.04012020-draft.pdf

[6] M. Jethanandani, S. Agarwal, L. Huang, and D. Blair, "YANG Data Model for Network Access Control Lists (ACLs)," RFC 8519, Mar. 2019. [Online]. Available: https://www.rfc-editor.org/info/rfc8519

[7] A. Anonymous, "Mudis - mud inspection system," 2021. [Online]. Available: https://github.com/ransh93/MUDIS

[8] B. Meyuhas, Shister, "Mud files dataset in different locations." 2021. [Online]. Available: https://github.com/barmey/IoT mud files locations

[9] V. Andalibi, E. Lear, D. Kim, and L. J. Camp, "On the analysis of mud-files' interactions, conflicts, and configuration requirements before deployment," arXiv preprint arXiv:2107.06372, 2021.

[10] N. I. o. S. NIST and Technology, "Mud-pd is a tool assist in the characterization of iot device network behavior and the creation and definition of appropriate mud files." [Online]. Available: https://github.com/usnistgov/MUD-PD

[11] A. M. Mandalari, R. Dubois, Daniel J.and Kolcun, M. T. Paracha, H. Haddadi, and D. Choffnes, "Blocking without breaking: Identification and mitigation of non-essential iot traffic," in Proc. of the Privacy Enhancing Technologies Symposium (PETS), 2021.

[12] "Mud maker tool," 2021. [Online]. Available: https://mudmaker.org/

[13] H. Guo and J. Heidemann, "Detecting iot devices in the internet," IEEE/ACM Transactions on Networking, vol. 28, no. 5, pp. 2323–2336, 2020.

[14] M. H. Mazhar and Z. Shafiq, "Characterizing smart home iot traffic in the wild," in 2020 IEEE/ACM Fifth International Conference on Internet-of-Things Design and Implementation (IoTDI). IEEE, 2020, pp. 203–215.

[15] G. Hu and K. Fukuda, "Toward detecting iot device traffic in transit networks," in 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC). IEEE, 2020, pp. 525– 530.

[16] S. A. Hamad, W. E. Zhang, Q. Z. Sheng, and S. Nepal, "Iot device identification via network-flow based fingerprinting and learning," in 2019 18th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/13th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE). IEEE, 2019, pp. 103–111.

[17] R. Perdisci, T. Papastergiou, O. Alrawi, and M. Antonakakis, "Iotfinder: Efficient large-scale identification of iot devices via passive dns traffic analysis," in 2020 IEEE European Symposium on Security and Privacy (EuroS&P). IEEE, 2020, pp. 474–489.

[18] Cisco, Jan 2018. [Online]. Available: https://www.cisco. com/c/en/us/td/docs/ios-xml/ios/sec data acl/configuration/xe-3s/ sec-data-acl-xe-3s-book/sec-cfg-fqdn-acl.html

[19] Y. Afek, A. Bremler-Barr, D. Hay, R. Goldschmidt, L. Shafir, G. Avraham, and A. Shalev, "Nfv-based iot security for home networks using mud," in NOMS 2020-2020 IEEE/IFIP Network Operations and Management Symposium. IEEE, 2020, pp. 1–9.

[20] NordVPN, "Leading vpn service. online security starts with a click." 2021. [Online]. Available: https://nordvpn.com/