REFINING MEASUREMENT OF INTERNET INFRASTRUCTURE PUBLIC HEALTH

Technical Report 22-03

Exploring the definitions, similarities, and differences between internet infrastructure, critical infrastructure, and internet infrastructure public health

CyberGreen Institute
March 2022

TABLE OF CONTENTS

1. EXECUTIVE SUMMARY

The effective measurement of the health of internet infrastructure is predicated on a precise definition. Without such precision, different groups are likely to measure different things. Many of the elements of internet infrastructure are hard to measure from within the internet. For example, the reliability of power is not visible to the computers plugged into electrical sockets.

This paper refines the definition of Internet Infrastructure Public Health (IIPH) by:

1. Separating out the "vital statistics" of IIPH. These are the population defining statistics which scope the idea of "a public" whose health is being measured
2. Enumerating criteria which can be applied to the selection of data to gather or use for IIPH.

2. BACKGROUND & RELATED WORK

There is closely related work in defining internet infrastructure and critical infrastructure. Nothing in this paper is intended to argue against extant definitions. There are two important groupings of infrastructure that relate to this work, and it is helpful to understand them. They are internet infrastructure and critical infrastructure.

Internet Infrastructure is defined by the Internet Infrastructure Coalition as a collective term for all hardware and software systems that are "responsible for hosting, storing, processing, and serving the information that makes up websites, applications, and content" (2019). There are other, broader definitions, such as one from ENISA which we discussed in the IIHMF report, and do not revisit here for space reasons (Lévy-Bencheton et al. 2015; Shostack & Kaeo 2021). Critical infrastructure broadly refers to those systems which are critical to the functioning of society, including power, water, and the like. The concept of critical infrastructure has become widespread, and there are many definitions of both critical infrastructure and critical internet (or information) infrastructure. The Fraunhofer Institute has cataloged many of these from around the world.

Some have criticized the definitions of critical infrastructure as being overly broad and imprecise (e.g. Jhangiani & Kennis 2022). We note that the U.S. Department of Homeland Security appears to define critical information infrastructure very broadly, possibly including every information system operated by a state or even local government.[1]

---

[1] The definition has a quite complex structure, including clauses separated by several and an "or". That "or" may be intended to mean A & (B OR C), that is A in combination with either B or C. It could also mean A or B or C, that is, any of A, B, C:

Critical information infrastructure (CII) is any physical or virtual information system that controls, processes, transmits, receives or stores electronic information in any form including data, voice or video that is: (1) Vital to the functioning of critical infrastructure; (2) So vital to the United States that the incapacity or destruction of such systems would have a debilitating impact on national security, national economic security, or national public health or safety; or (3) Owned or operated by or on behalf of a State, local, tribal, or territorial government entity. (Adapted from DHS, 2011 by the CIPR project of the Fraunhofer Institute.)].

Both internet infrastructure and critical infrastructure thinking relate to the public health of internet infrastructure, as we discuss in the next section.

3. INTERNET INFRASTRUCTURE PUBLIC HEALTH (IIPH)

Cyber Public Health adopts and adapts public health thinking to the cause of preventing cyber attacks and the promotion of a more secure and resilient internet for all through informed choices of society, organizations and communities.

We define Internet Infrastructure Public Health (IIPH) as using the tools of cyber public health to measure internet infrastructure. This means that IIPH must be focused on a subset of internet infrastructure given the immediate limitation of "measurability".

We can measure 3 aspects of IIPH:

- Vital Statistics: How many systems are included in the internet infrastructure?
- Security of Systems: How securely configured and managed are the systems which make up internet infrastructure? (See important notes about the practicality of these measurements in the "Components of IIPH" section below.)
- Harm to Others: What decisions are organizations making that have an impact on the safety of others? For example, not filtering source addresses allows spoofed IP packets to propagate, not applying BGP security measures weakens the global routing infrastructure. Neither of these particularly harm the ISPs which fail to do them.

The qualifier of "health" is amongst the trickiest parts of this project. There are measures in classic public health, such as COVID-19 infections, which are both personal medical information and, at scale, public health information. In earlier work (Shostack & Kaeo 2021), we excluded Security of Systems as an aspect of IIPH. That artificial separation of "harm to self" from "harm to others" was an accessible means of distinguishing cyber public health from enterprise or personal security. This was useful, and more recent work including a structured analysis of public health, causes us to include Security of Systems (Shostack 2022).

4. COMPONENTS OF IIPH

We could measure many things, and we apply a principle-oriented approach to selecting targets of measurement. We are in the process of defining and discovering and refining the principles which will matter, and our understanding currently includes:

- Illustrative: does it show something about infrastructure?
- Practical considerations
- A set of selection practicality considerations
- Transparency considerations

Illustrative:

- Cybersecurity is awash in data graphics, soundbites and data points that are hard to interpret;
- Most of this data is about people or firms, rather than infrastructure (internet, critical or otherwise);
- Most of this data attempts to sell products or services, rather than assess or inform policy.

Selection Practicality:

- Can we gather data accurately and consistently?
- Is the data quantifiable?
- Can the data be traced back to a system we already report on?
- Redundancy: Is the data very similar to data we already collect?
- Is anyone doing it? There are proposals, such as certificate pinning and DNSSec, which are rare in practice. The Cyber Belief Model[2] may be in play, or the downsides of the technology may be great enough that no one does it;
- Legal concerns: can we gather data without permission and without substantial legal risk?

Transparency Considerations:

- Quantitative transparency: There are infographics, charts, and other sources which may even have quantitatively labeled axes (but not always). We look for data whose source will provide us with numbers, rather than some derivative work;
- Methodological transparency: There is a great deal of data gathered where it is not clear how the data is gathered, what population is studied, what choices the researchers have made, or what they believe are the limits of the data. We look for data whose source provides us with methodological information.

Temporal Considerations:

- Is the data being gathered regularly by an organization which expects to continue over time?
    - If not, is the data substitutable or replaceable, or does it depend on that organization's unusual data gathering perspective or capabilities?
- What is the rate of change? A measure which fluctuates wildly day to day may be hard to interpret, one that never changes may be irrelevant.

Vital statistics: There are a few components of IIPH whose measurement begins with counting, like the vital statistics of births and deaths. For example, how many routers or DNS servers exist? How frequently do the numbers change? Without such population statistics, our assurance in the data we gather is inevitably limited, and our ability to assess prevalence or incidence is curtailed.

System Properties: We look for information about the properties of specific systems, which can be a specific host, a service (such as "the DNS for JPCERT") or a network or Autonomous System.

[2] A Health Belief Model is in widespread use in public health, and we are exploring a parallel model Cyber Belief Model), and expect to publish on it this year.

An earlier list of considerations is in (Shostack & Kaeo, 2021). That list is longer; the one above is better organized.

VISUALIZING THE RELATIONSHIP BETWEEN INTERNET INFRASTRUCTURE, CRITICAL INFRASTRUCTURE, AND IIPH

It may be helpful to visualize the relationship between IIPH, internet infrastructure and critical infrastructure. We use a set of Venn Diagrams here; they are drawn to show the inter-relationships, not to imply anything about relative scales. This paper uses the conceptualization shown in Figure 1, and we present the others to help readers understand other approaches that could have been taken.

There are elements of both internet infrastructure and critical infrastructure which are outside our assessment of health. For example, food supply is absolutely critical infrastructure, and outside our focus. Internet infrastructure includes capacity, and assessments of infrastructure may include elements of geographic coverage, affordability, or the like which are similarly outside our focus.

IIPH may be a subset of both internet infrastructure, as defined by the Internet Infrastructure Coalition and a subset of critical infrastructure (in most definitions of that term). There are also elements of IIPH which relate to Cyber Public Health, outside these Infrastructures. For example, public health concerns itself with vital statistics of population, and cyber public health concerns itself with technical populations. Infrastructure is a subset of those populations and is not shown here.

One way to conceptualize this is shown in Figure 1 where IIPH is entirely contained within internet infrastructure, and contains elements of critical infrastructure. The "mirror" set of Critical Infrastructure Public Health, which excludes some internet infrastructure, is excluded from this project because this project is focused on internet infrastructure. We do not mean to imply that Critical Infrastructure Public Health is not something we could consider.

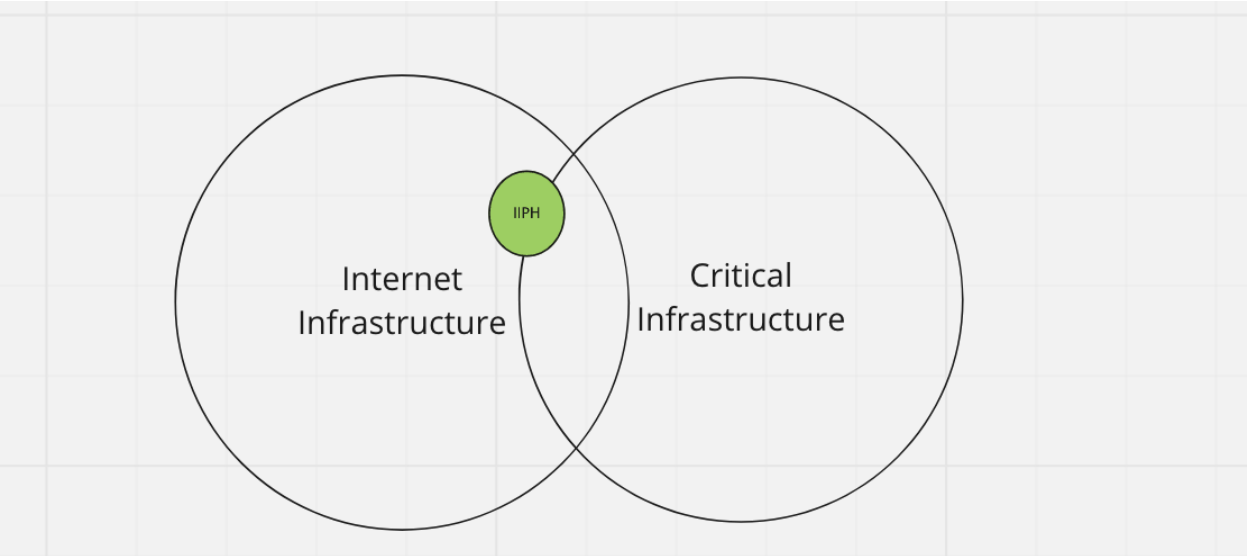

Figure 1: Elements of critical infrastructure

Another way to think of IIPH is that it only concerns itself with that internet infrastructure which is critical infrastructure. This is shown in Figure 2:

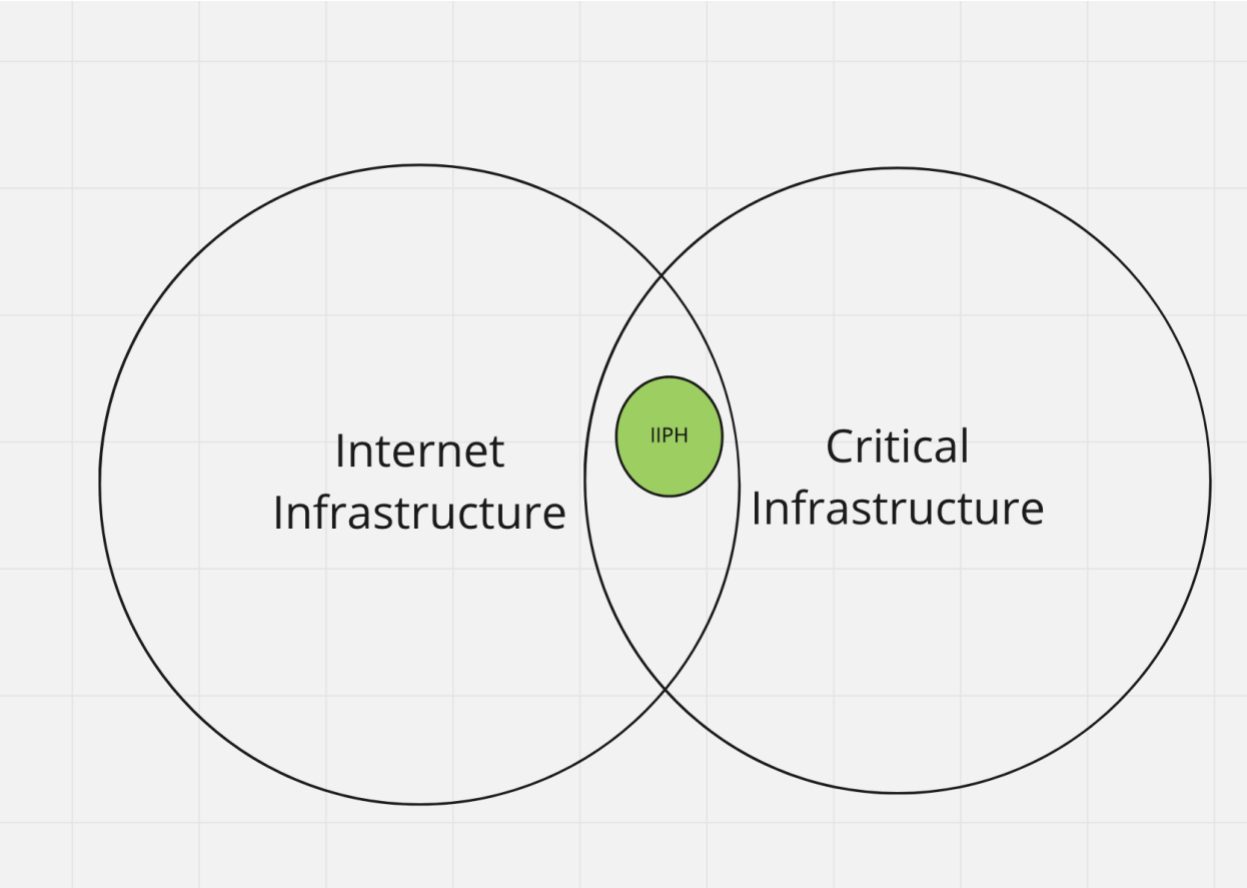

Figure 2: The subset definition

It is also possible that there are elements of Internet Infrastructure Public Health which are outside the definition of either internet infrastructure or critical infrastructure. This conceptualization is shown in Figure 3. For example, if my neighborhood internet is down, that may be too small to impact on either, but it may impact on my trust in the system.

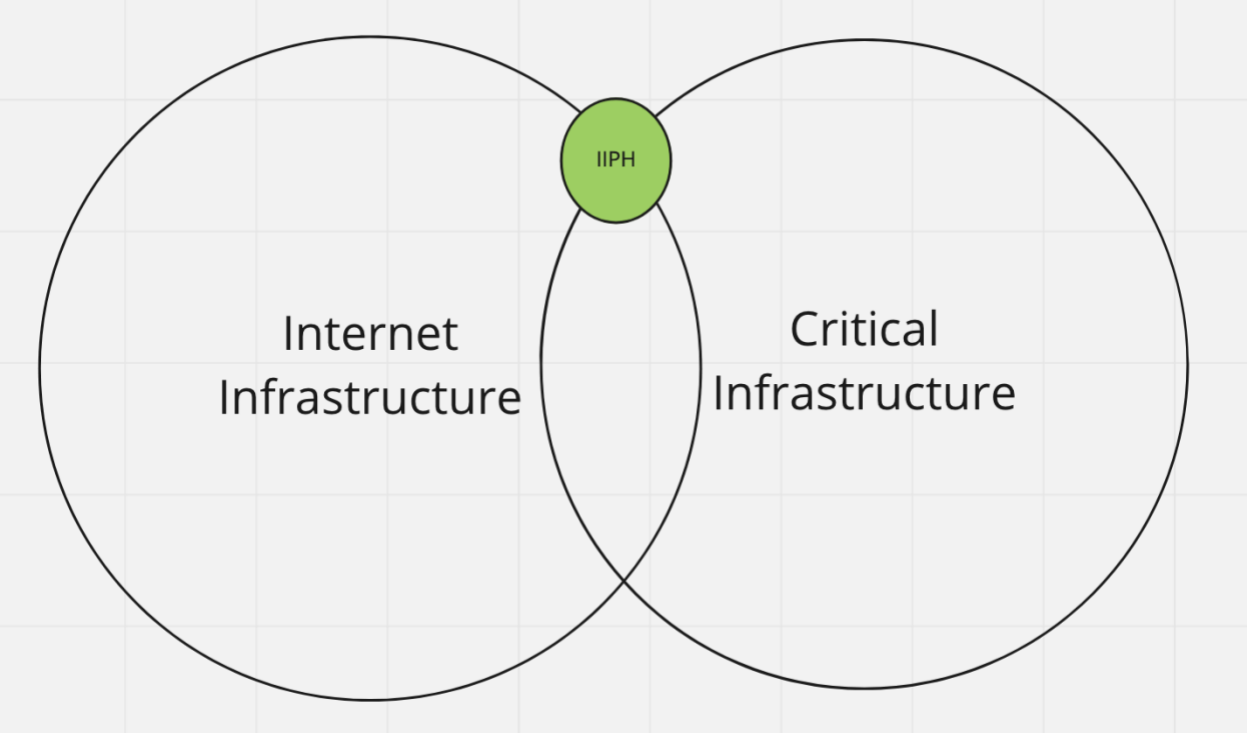

Figure 3: The overlap definition

5. CONCLUSION

This technical report refines the definition of Internet Infrastructure Public Health as a follow-up to (Shostack & Kaeo, 2021). It defines the relationship between closely related infrastructures, and a set of considerations for what to measure.

The illustration of the relationship between internet infrastructure, critical infrastructure and public health will help readers better place this work, and aid in continuing refinement.

The considerations presented are an important building block. Collaborators frequently urge us to look at this measurement or that, and being able to ask specific, principle-guided questions of what system is covered, and assess if the statistic would be illustrative, practically feasible to gather, and available in useful timescales will help focus our work and improve understanding.

The public health metaphor offers an exciting and different possibility for advancing cybersecurity, and this paper is another step towards advancing that goal.

REFERENCES

CIPedia contributors, "Critical Information Infrastructure," *CIPedia, ,* https://websites.fraunhofer.de/CIPedia/index.php?title=Critical_Information_Infrastructure&oldid=12224 (accessed March 2022).

Department of Homeland Security (DHS), Blueprint for a Secure Cyber Future § (2011). https://www.dhs.gov/xlibrary/assets/nppd/blueprint-for-a-secure-cyber-future.pdf.

Internet Infrastructure Coalition. "What Is the Internet's Infrastructure?" Internet Infrastructure Coalition, February 6, 2019. https://www.i2coalition.com/what-is-the-internets-infrastructure-video/.

Jhangiani, Tasha, and Graham Kennis. "Protecting the Critical of Critical: What Is Systemically Important Critical Infrastructure?" Web log. Lawfare (blog), June 15, 2021. https://www.lawfareblog.com/protecting-critical-critical-what-systemically-important-critical-infrastructure.

Lévy-Bencheton, Cédric, Louis Marinos, Rossella Mattioli, Thomas King, Christoph Dietzel, and Jan Stumpf. "Threat Landscape and Good Practice Guide for Internet Infrastructure." ENISA, January 2015. https://www.enisa.europa.eu/publications/iitl/view/++widget++form.widgets.fullReport/@@download/Threat+Landscape+and+Good+Practice+Guide+for+Internet+Infrastructure.pdf

Shostack, Adam, and Merike Kaeo. Rep. Internet Infrastructure Health Metrics Framework (IIHMF). CyberGreen Institute, 2021. https://www.cybergreen.net/img/medialibrary/IIHMF.pdf.

Shostack, Adam. Tech. Vital Statistics in Cyber Public Health. CyberGreen Institute, 2022.

ACKNOWLEDGEMENTS

This report was written by Adam Shostack for the CyberGreen Institute.