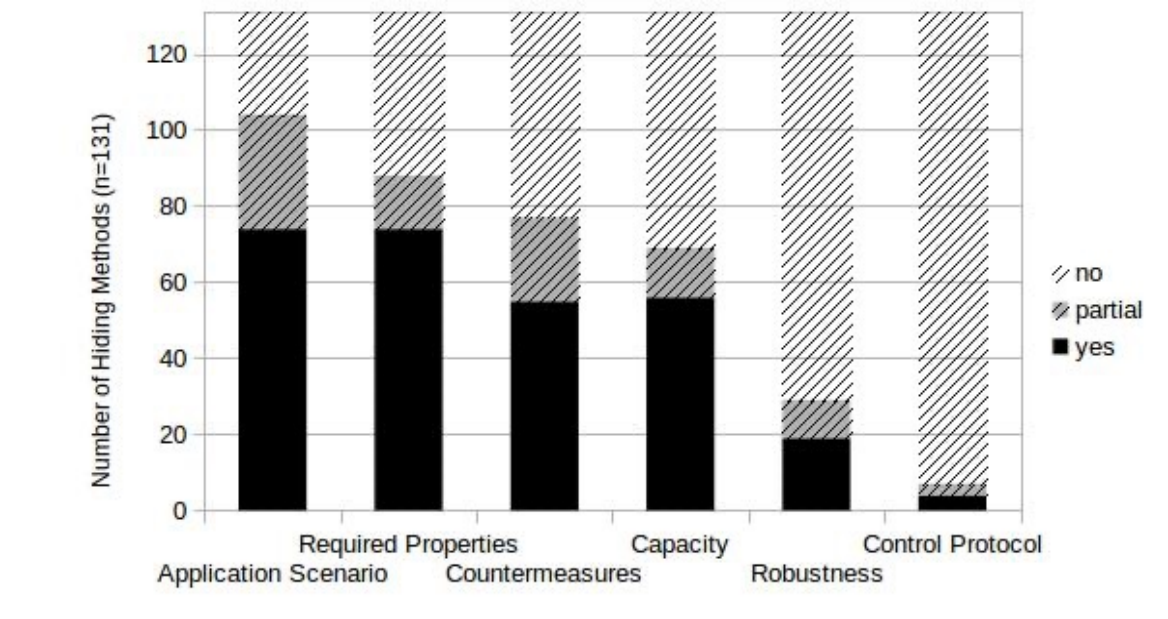
INTERNET CENSORSHIP

CH. 9: DESCRIBING HIDING METHODS PLUS SOME NOTES ON EXPERIMENTAL SETUPS

Prof. Dr. Steffen Wendzel
https://www.wendzel.de

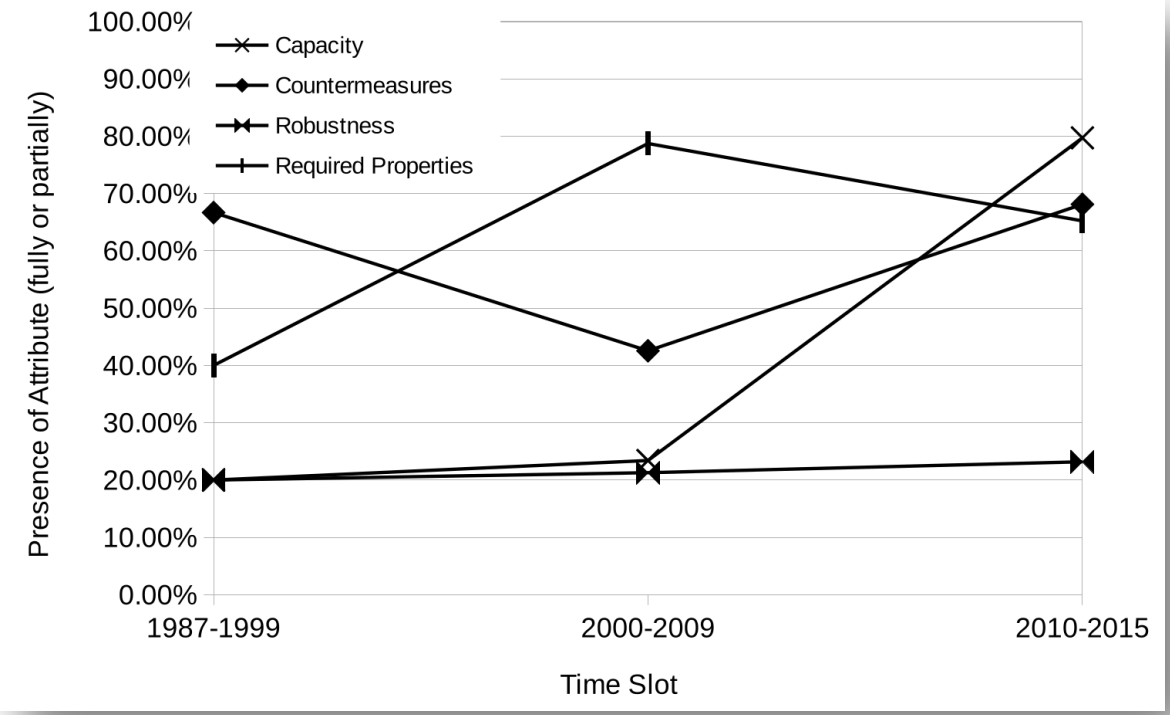
Analysis of 131 Hiding Techniques [1]

- Descriptions of hiding techniques in scientific papers highly vary, rendering it difficult to compare them:



[1] S. Wendzel, W. Mazurczyk, S. Zander: Unified Description for Network Information Hiding Methods, in: Journal of Universal Computer Science, 2016.

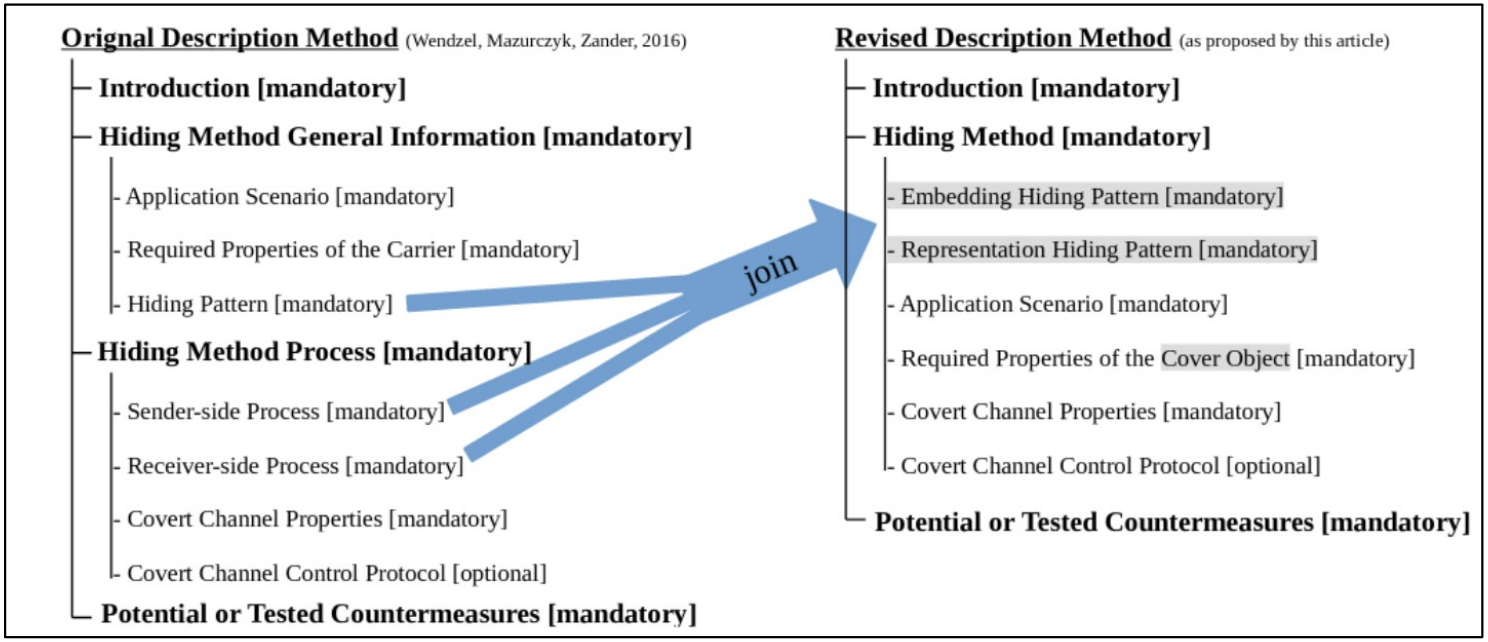
Analysis of 131 Hiding Techniques [1]



[1] S. Wendzel, W. Mazurczyk, S. Zander: Unified Description for Network Information Hiding Methods, in: Journal of Universal Computer Science, 2016.

Describing Hiding Methods Using Patterns [1]

- We proposed a method to unify the descriptions within new publications. Our method is simply called a **unified description method**.
- Detailed description of the attributes + examples can be found in the paper.



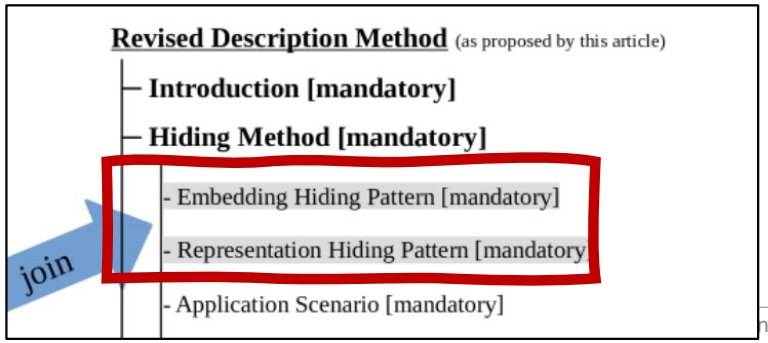
[1] S. Wendzel et al.: Generic Taxonomy for Steganography Methods, ACM Comp. Surv., 2025 (see supplemental material of the paper).

Only “Embedding Pattern” and “Representation Pattern”? [1]

- No, we should enhance the name of a pattern with some more specific terms.

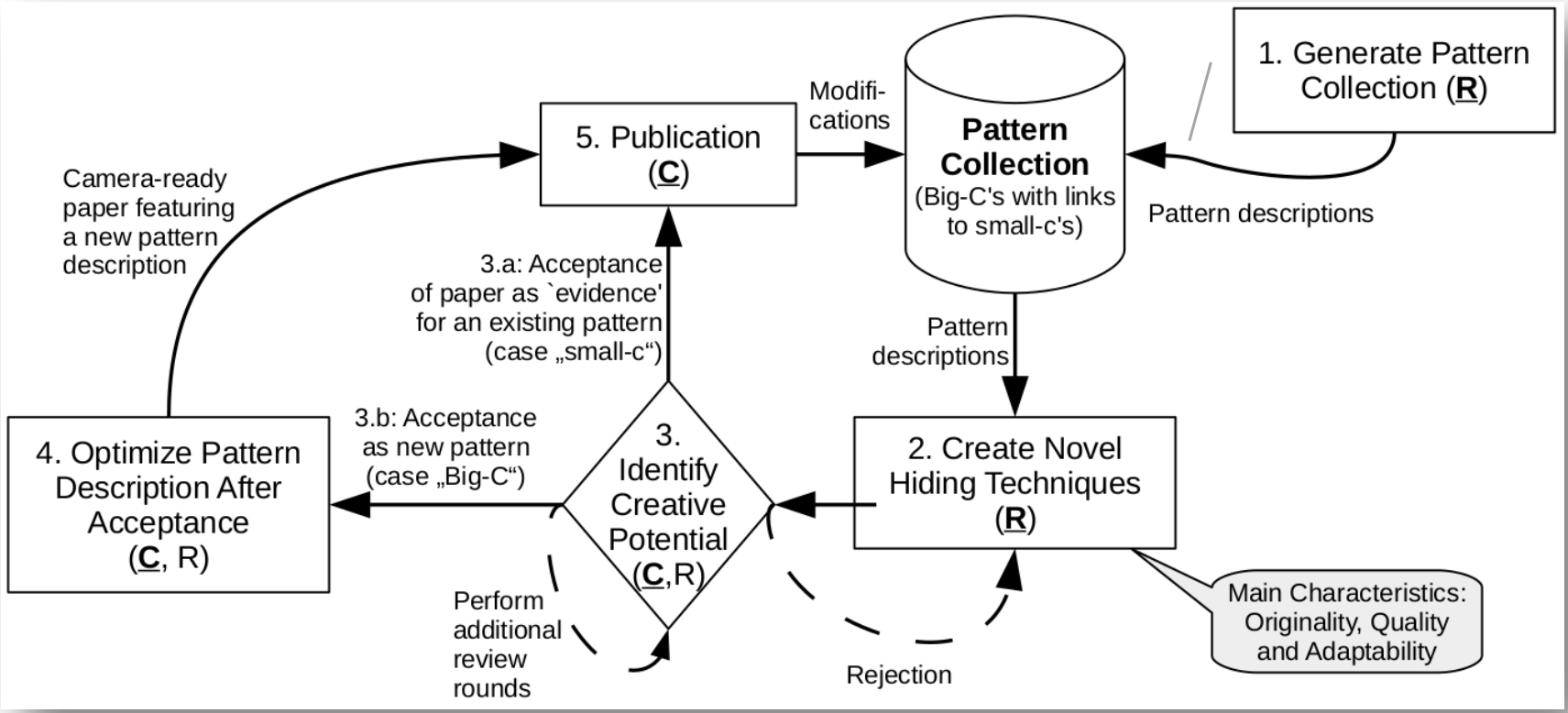
  Interactive Tool: https://patterns.omi.uni-ulm.de/desrcovert/

Locality		**Directness**		**Activeness** [Zander‘10 / Lamshöft & Dittmann’20 / Wendzel et al.‘25]		**Levels**		**Reference-temporality** [Wendzel et al.’25]			*	**Hiding Pattern** [Wendzel et al.’25]
- (non)	distributed	- (direct)	indirect	active	not (purely) active	- (single)	multi-level	history	- (present)	future		
	Pattern of [Wendzel et al.’15, Mazurczyk et al.’18] (pattern comb., pattern hopping, pattern variation)		Pattern of [Schmidbauer et al. ’22] (redirector, dead proxy, drop)		semi-active, semi-passive, passive, fully-passive							



[1] S. Wendzel et al.: *Combining Different Existing Methods for Describing Steganography Hiding Methods*, in Proc. ARES 2025 Workshops, 2025.

Patterns as a Tool to Prevent Scientific Re-Inventions [1]

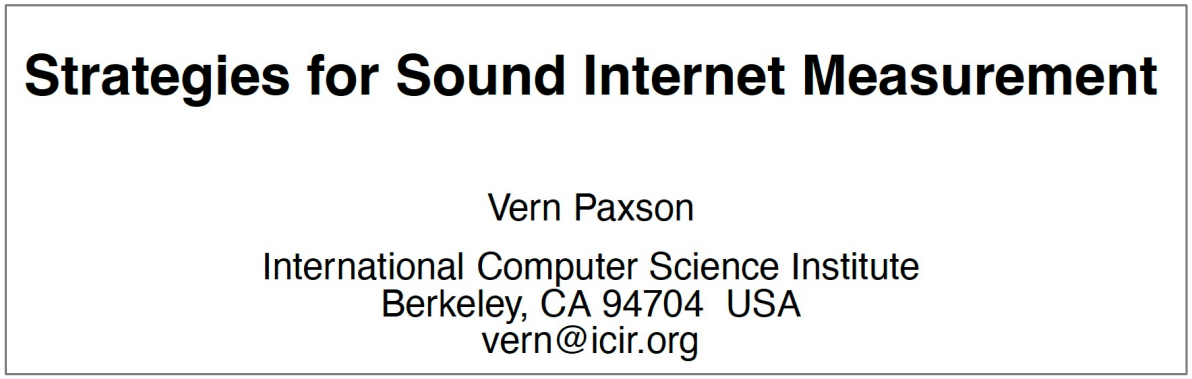


[1] S. Wendzel and C. Palmer: *Creativity in Mind: Evaluating and Maintaining Advances in Network Steganographic Research*, J.UCS, Vol. 21, 2015.

Notes on General Measurement Studies

(based on/partially copied from: Paxson: Strategies for Sound Internet Measurement, Proc. IMC, 2004)

- Paxon summarizes observations that aim to „help students … avoid some of the pitfalls that practitioners have come to appreciate over time“.



Strategies for Sound Internet Measurement

Vern Paxson

International Computer Science Institute
Berkeley, CA 94704 USA
vern@icir.org

Notes on General Measurement Studies

(based on/partially copied from: Paxson: Strategies for Sound Internet Measurement, Proc. IMC, 2004)

- **Dealing with Errors and Imperfections**
  - **Precision** ("the maximum exactness that a tool's design permits"):
    - Example: a clock reports time in units of 1 µsec, so it cannot record any distinction in time finer than that. However, due to OS-implementation aspects, the advancement of the clock might be much coarser (e.g., advancing only every 10 µsec).
      - Thus: if exact timing is relevant, reports should include an indication of a tool's precision.
  - **Meta-data** associated with measurement traces:
    - Preserving the information during the course of analysis.
    - Example: network packet trace formats do not allow for meta-data to be embedded.
      - Early examples: `ipsumdump` produces configurable ASCII output from traces that can feature annotations (e.g., details about the host and the options used to record the trace).

Notes on General Measurement Studies

(based on/partially copied from: Paxson: Strategies for Sound Internet Measurement, Proc. IMC, 2004)

- **Dealing with Errors and Imperfections**
  - **Accuracy** (measurements are an abstraction of a phenomenon, but: **how well does the measurement match the actual phenomenon?**)
    - Example: packet filters record traffic matching a filter expression. However, they might fail to record *all* packets matching the used filter expression ("packet filter drop").
      - Can be caused by failure of the software (`wireshark`/`tcpdump`/…) or due to a failure of the filter to keep up with the rate at which packets are provided to the filter by the tap.
      - Measurement tools often do not report all relevant kinds of such failures.
        - Example: `tcpdump` provides an end-of-run summary of **total drops**. However, this **does not necessarily include drops by the tap itself**.
          - Moreover: meta-data is not included in the trace file, so meta-data is separated from the recording.
    - Example: Unlike precision, clock inaccuracy can be caused by failure of synchronization to true time or clock jumps.
      - One might use GPS clock synchronization instead of NTP. However, the OS kernel might still provide incorrect clock settings due to packet buffering -> incorrect packet timestamps.

Notes on General Measurement Studies

(based on/partially copied from: Paxson: Strategies for Sound Internet Measurement, Proc. IMC, 2004)

- **Dealing with Errors and Imperfections**
  - **Misconception** ("errors in equating what we are *actually* measuring with what we *wish* to measure"). Examples:
    - **TCP packet loss**: measured by counting retransmitted packets. Risk: overlooking that packets might be retransmitted unnecessarily or retransmission packets being replicated by the network.
    - **Measuring web transfer times** without accounting for hidden proxies.
      - See Paxson's paper, p. 265
    - **Quantifying TCP throughput** using transfers with large socket buffers and modest transfer sizes.
      - Measuring on the application level (e.g., "how long it takes to issue all of the `write` system calls") only measures how long it takes to fill the kernel buffer.
    - "**Computing the distribution of TCP connection sizes** by capturing SYN and FIN packets and using the difference between their seq. no.".
      - Maybe failing to recognize that the very largest connections […] might often already be underway when we start tracing, or have not terminated when we finish".
    - **Vantage points** (as also used for Internet censorship measurement): location "can significantly skew the interpretation of the measurement". -> already discussed during previous lectures!

Notes on General Measurement Studies

(based on/partially copied from: Paxson: Strategies for Sound Internet Measurement, Proc. IMC, 2004)

- **Dealing with Errors and Imperfections**
  - **Misconception** (cont.). Examples:
    - We might **falsely assume a traffic recording to be representative** but it might not necessarily be! Some attributes might be typical for specific network conditions (e.g., ARP request behavior) but others (e.g., median download item size) might not.
      - Use datasets of different locations and settings as well as recorded at different times
        - Example: try accessing a plethora of websites from different vantage point locations.
      - Seek out peer-review at an early stage of your research project

Notes on General Measurement Studies

(based on/partially copied from: Paxson: Strategies for Sound Internet Measurement, Proc. IMC, 2004)

- **Dealing with Errors and Imperfections**
  - **Calibration:**
    - **Examining outlies** (unusually low/high values) and **spikes** (values that repeat very often). Use these as a clue to double check for measurement errors, analysis errors, or misconceptions.
      - Examples:
        - **Measurement error:** (1) figuring out that that some RTT values are physically infeasible.
        - **Analysis error:** (1) RTT computed the wrong way due to a mismatch in associating the outbound packet with the wrong reply; (2) computing connection sizes using TCP SYN/FIN/RST packets but {having a not considered integer overflows}. // paper provides a slightly different reason.
        - **Misconception error:** See Paxson's paper, p. 266.

Notes on General Measurement Studies

(based on/partially copied from: Paxson: Strategies for Sound Internet Measurement, Proc. IMC, 2004)

- **Dealing with Errors and Imperfections**
  - **Calibration (cont.):**
    - **Self-consistency checks**:
      - Example: **Determining whether a trace includes all of the traffic associated with a given connection**, or whether some of the traffic is missing due to a measurement problem, such as a packet filter drop -> take advantage of TCP reliability:
        - TCP receiver should never send an acknowledgement for data it has not received, and TCP acknowledgements are cumulative. Thus, we can at least check if a receiver did indeed receive all the segments (at least once, because it might still include retransmissions due to packet loss!).

Notes on General Measurement Studies

(based on/partially copied from: Paxson: Strategies for Sound Internet Measurement, Proc. IMC, 2004)

- **Dealing with Errors and Imperfections**
  - **Calibration (cont.):**
    - **Comparing multiple measurements**:
      - Measuring the same phenomenon in different ways and comparing the results.
      - Example: run two separate packet monitors to see if their recordings match or if they face differences, e.g., different packets not being recorded. E.g., one monitor *downstream* and one monitor *upstream*. (Additional reasons might come into the game here, e.g., downstream packet taking a different route than an upstream packet).

Notes on General Measurement Studies

(based on/partially copied from: Paxson: Strategies for Sound Internet Measurement, Proc. IMC, 2004)

- **Dealing with Errors and Imperfections**
  - **Calibration (cont.):**
    - **Comparing multiple measurements (cont.)**:
      - Fig. 1 shows one-way transit times (timestamp of receiver upon arrival - timestamp of sender upon transmission)
        - At around 750msec: sudden downward shift
        - Could be caused by routing change [a network event] or clock adjustment at either sender or receiver [a measurement glitch!].
      - Fig. 2 shows the same measurement but including data of a reverse-path measurement, too (hollow squares), computed in the same way.
        - Equal-but-opposite character of the jump in two directions tells us that it is *likely* caused by a clock jump as it would match ”what we would expect from a clock jump”. -> a case of calibration rooted in *plausibility* rather than direct comparison.

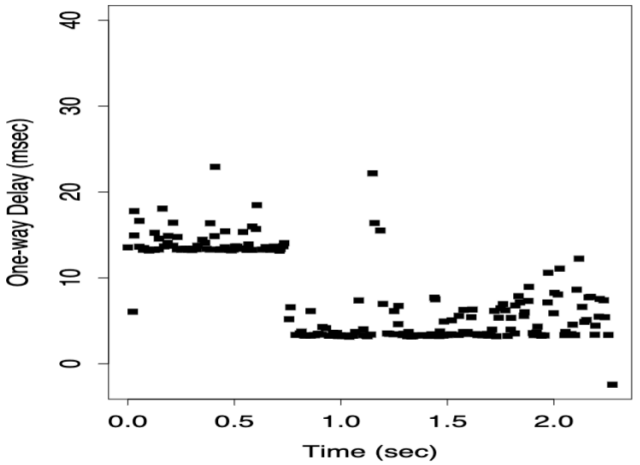


Figure 1: One-way transit time step that could be due to either a routing change or a clock adjustment.

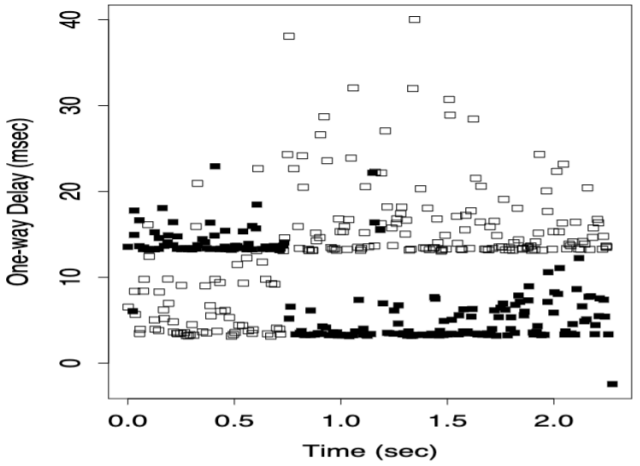


Figure 2: Incorporating additional measurements resolves the change as due to a clock adjustment.

Notes on General Measurement Studies

(based on/partially copied from: Paxson: Strategies for Sound Internet Measurement, Proc. IMC, 2004)

- **Dealing with Errors and Imperfections**
  - **Calibration (cont.):**
    - **Evaluating synthetic data:**
      - Testing with artificially crafted data that matches expected data to see whether our analysis is sound. For instance, hand-crafted web server logs, followed by hand-crafted modifications.
      - Conducting Monte Carlo simulations (providing data that follows some desired statistical distribution and see whether our results are sound). For instance, this could be applied to the Berk et al. metric discussed in earlier chapters of this class.

Notes on General Measurement Studies

(based on/partially copied from: Paxson: Strategies for Sound Internet Measurement, Proc. IMC, 2004)

- **Dealing with Errors and Imperfections**
  - **Large-volume Data:**
    - System limitations, such as disk space, max. file size, memory limitations, max. number of files on the filesystem etc.
      - Additionally relevant when creating plots:
        - Simple approach is to strip out redundant points during visualization.

- **FAIR (not part of Paxson's paper but copied from https://data.research.cornell.edu/data-management/sharing/fair/):**
  - **Findable:** data and metadata are online and openly searchable with a persistent link that is uniquely attached to each specific dataset.
  - **Accessible:** data and metadata are retrievable in machine-actionable form, with downloading options clearly described (including any needed authentication).
  - **Interoperable:** data and metadata are consistently structured and described, both syntactically and semantically, so that algorithms can parse and ensure like data are accurately compared to like.
  - **Reusable:** data and metadata are sufficiently annotated so machine and human users can determine fit-for-purpose in the context of their analysis.

Notes on General Measurement Studies

(based on/partially copied from: Paxson: Strategies for Sound Internet Measurement, Proc. IMC, 2004)

- **Reproducibility:**
  - Partially already discussed. However, some more notes from Paxson's paper:
    - Months after experiments have been conducted and research work has been submitted to review, reviewers ask to conduct experiments in a slightly different way.
    - Do you remember what you did? Do your experiments still work with (updated) software?
    - Describe all experimental characteristics and steps in detail (also to aid your own future work!). Keep a notebook, use a version control system (we use Git + Overleaf with tracking changes).
      - -> lets you also undo certain things where you have taken the wrong path.
    - Paxson recommends a single "master script" that "builds all analysis results from the raw data" and "maintains all intermediary, reduced forms of the data as explicitly ephemeral."
  - Make your data publicly available.
  - Describe your data in a unified manner (see the "Unified Description Method" that was already introduced in this class as an easy-to-apply example).