Limitations of Traffic Analysis at Large Scale

Timothy J. Shimeall
CERT/NetSA
FloCon 2013

Notices

NO WARRANTY

THIS MATERIAL OF CARNEGIE MELLON UNIVERSITY AND ITS SOFTWARE ENGINEERING INSTITUTE IS FURNISHED ON AN “AS-IS" BASIS. CARNEGIE MELLON UNIVERSITY MAKES NO WARRANTIES OF ANY KIND, EITHER EXPRESSED OR IMPLIED, AS TO ANY MATTER INCLUDING, BUT NOT LIMITED TO, WARRANTY OF FITNESS FOR PURPOSE OR MERCHANTABILITY, EXCLUSIVITY, OR RESULTS OBTAINED FROM USE OF THE MATERIAL. CARNEGIE MELLON UNIVERSITY DOES NOT MAKE ANY WARRANTY OF ANY KIND WITH RESPECT TO FREEDOM FROM PATENT, TRADEMARK, OR COPYRIGHT INFRINGEMENT.

Use of any trademarks in this presentation is not intended in any way to infringe on the rights of the trademark holder.

This Presentation may be reproduced in its entirety, without modification, and freely distributed in written or electronic form without requesting formal permission. Permission is required for any other use. Requests for permission should be directed to the Software Engineering Institute at permission@sei.cmu.edu.

This work was created in the performance of Federal Government Contract Number FA8721-05-C-0003 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center. The Government of the United States has a royalty-free government-purpose license to use, duplicate, or disclose the work, in whole or in part and in any manner, and to have or permit others to do so, for government purposes pursuant to the copyright license under the clause at 252.227-7013.

What We Will Cover

Overarching questions

What will we never know?

Analytical Limitations

Overarching Questions

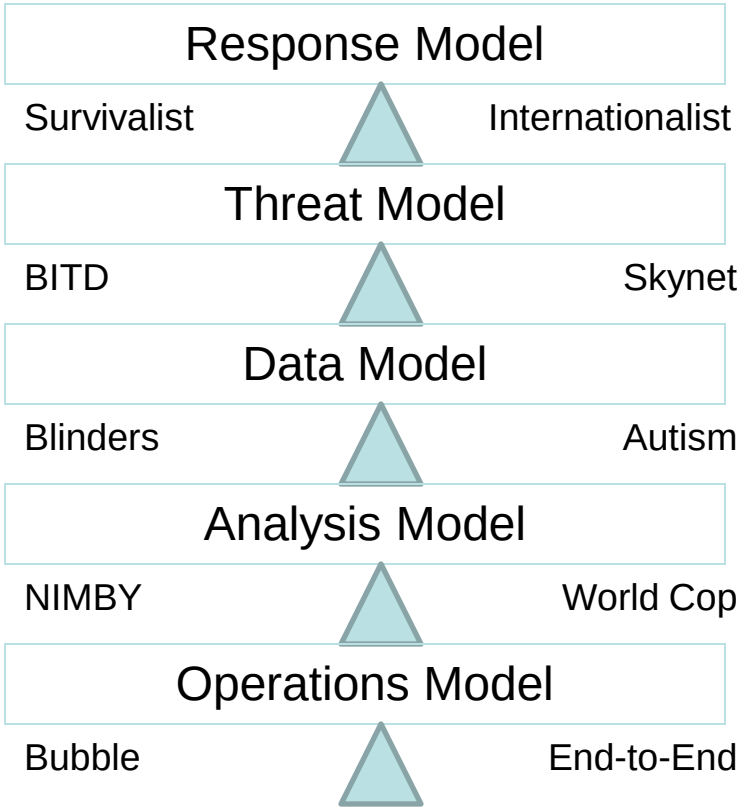
How do I know what I’m looking for?

How do I know why I’m looking?

How do I know where to look?

How do I know when it’s found?

Traffic Balance



What will you never know?

Total ground truth on a network of any size

How often is bad considered good? (false negative)

What is the next attack?

Why did they attack you?

What are your competitors seeing?

Inherently Partial Data

Technology shifts

Attacker actions

Defender actions

Managerial decisions

Network bandwidth

Correlation and Causation

Baseline in dynamic environment

Correlation vs. Causation

Implications

- Need to be cautious in kinds of conclusions
- Consider strategies for dealing with analysis gone wrong

Indication and Proof

Indication: There is reason to believe

Proof: There is no other logically defensible explanation

How much confidence do you need?

Cost of false positive?

Cost of false negative?

Conclusions

Many failure modes

Many challenges

Topic of continuing interest