A Bibliography on Network Analytics

Professor Salvatore J Stolfo
Department of Computer Science
Columbia University
New York, NY 10027-7003
sal@cs.columbia.edu
www.cs.columbia.edu/~sal

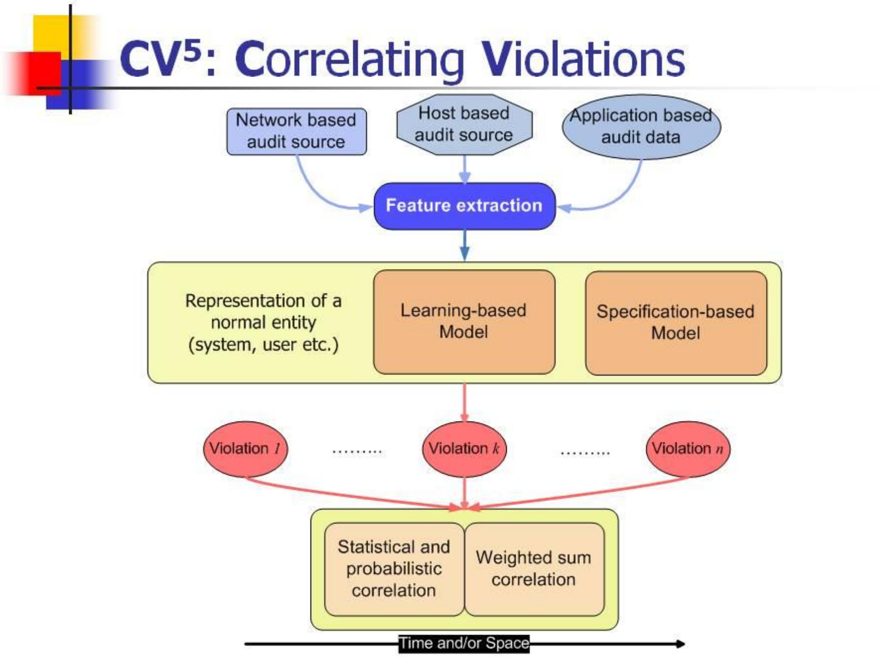
The ***CV***^{5} framework provides a convenient way of organizing a vast literature on Network traffic and Host analysis for various security purposes, including cyber defense of critical infrastructure and identification of advanced slow and stealthy attack. CV^5 refers to the

Correlation of sensor alerts that indicate *Violations* of:

- *Velocity* (Avg/Max Frequency of distinct events)
- *Volume* (Avg/Number of events in unit time)
- *Values* (Consistency/probability of data)
- *Vertices* (Connections/relationships/cliques)

The framework was first introduced in
Salvatore J. Stolfo, Shlomo Hershkop, Chia-Wei Hu, Wei-Jen Li, Olivier Nimeskern, Ke Wang "Behavior-based Modeling and its Application to Email Analysis" *ACM Transactions on Internet Technology (TOIT)* , Feb 2006.
http://sneakers.cs.columbia.edu/ids/publications/TOIT-EMT.pdf

Figure 1 provides a general view of network- and host-based audit and analysis using this framework.


CV⁵: Correlating Violations
Network based audit source
Host based audit source
Application based audit data
Feature extraction
Representation of a normal entity (system, user etc.)
Learning-based Model
Specification-based Model
Violation 1
....
Violation k
....
Violation n
Statistical and probabilistic correlation
Weighted sum correlation
Time and/or Space

Many of the papers selected for this bibliography appear in the literature in recent years; citations therein point to many prior papers.

Velocity

Slow Velocity of Probes Seth Robertson, Eric V. Siegel, Matt Miller, and Salvatore J. Stolfo. ``Surveillance Detection in High Bandwidth Environments." *In Proceedings of the 2003 DARPA DISCEX III Conference*. April, 2003.
http://sneakers.cs.columbia.edu/ids/publications/SD-DiscexIII.pdf

Velocity of packet arrivals of a flow in fired-length time intervals C.-M. Cheng, H. Kung, and K.-S. Tan, "Use of spectral analysis in defense against DoS attacks," in Proceedings of the IEEE GLOBECOM, Taipei, China, 2002

Velocity of packet arrivals of a flow in fired-length time intervals Alefiya Hussain, John Heidemann, and Christos Papadopoulos. "A Framework for Classifying Denial of Service Attacks". In Proceedings of ACM SIGCOMM 2003

Velocity of packets at each router and use collaborative anomaly detection among routers in order to detect Shrew DDos attacks, reduce the false positives this way Collaborative Detection Y. Chen and K. Hwang, "Collaborative Detection and Filtering of Shrew DDoS Attacks using Spectral Analysis," Journal of Parallel and Distributed Computing, Special Issue on Security in Grids and Distributed Systems, Vol. 66, Issue 9, September 2006

Volumetric

Volume of MIB variables M Thottan, C Ji, "Anomaly Detection in IP Networks", IEEE TRANSACTIONS ON SIGNAL PROCESSING, 2003

Volume of emails Gupta A. and Sekar R. "An approach for detecting self-propagating email using anomaly detection", in RAID, 2003.

Network bandwidth allocated to a flow R. H. K. Chen-Nee Chuah, Lakshminarayanan Subramanian. Dcap: Detecting misbehaving flows via collaborative aggregate policing. In *SIGCOMM Computer Communication Review*, volume 33, October 2003.

Volume of traffic A.Lakhina, M.Crovella, and C.Diot, "Diagnosing Network-Wide Traffic Anomalies," in Proc. of ACM SIGCOMM, 2004.

Volume of connections to a un-serviced port or to a vulnerable port: Senthilkumar G. Cheetancheri, John Mark Agosta, Denver H. Dash, Karl N. Levitt, Jeff Rowe, Eve M. Schooler, "A Distributed Host-Based Worm Detection System", Proceedings of the ACM SIGCOMM Workshop on Large Scale Attack Defense (LSAD06).

Haining Wang, Danlu Zhang, and Kang G. Shin. Detecting syn flooding attacks. In Proceedings of INFOCOM 2002, New York City, New York, June 2002.
http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.11.5580

A.Lakhina, M.Crovella, and C.Diot, "Diagnosing Network-Wide Traffic Anomalies," in Proc. of ACM SIGCOMM, 2004
http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.1.1838

M Thottan, C Ji, "Anomaly Detection in IP Networks", IEEE TRANSACTIONS ON SIGNAL PROCESSING, 2003
http://www.google.com/url?sa=t&source=web&ct=res&cd=1&url=http%3A%2F%2Fusers.ece.gatech.edu%2F~jic%2Fsig03.pdf&ei=Ei45SsemHouJtgfgktTiDA&usg=AFQjCNEZAzw1QIIE8FcPKmE8o_wHY4jXxw

C.-M. Cheng, H. Kung, and K.-S. Tan, "Use of spectral analysis in defense against DoS attacks," in Proceedings of the IEEE GLOBECOM, Taipei, China, 2002
http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.66.1487

Alefiya Hussain, John Heidemann, and Christos Papadopoulos. "A Framework for Classifying Denial of Service Attacks". In Proceedings of ACM SIGCOMM 2003 -#261
http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.127.8035

Gupta A. and Sekar R. "An approach for detecting self-propagating email using anomaly detection", in RAID, 2003
http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.6.2370

Y. Chen and K. Hwang, "Collaborative Detection and Filtering of Shrew DDoS Attacks using Spectral Analysis," Journal of Parallel and Distributed Computing, Special Issue on Security in Grids and Distributed Systems, Vol. 66, Issue 9, September 2006
http://portal.acm.org/citation.cfm?id=1232116

R. H. K. Chen-Nee Chuah, Lakshminarayanan Subramanian. Dcap: Detecting misbehaving flows via collaborative aggregate policing. In /*SIGCOMM Computer Communication Review*/, volume 33, October 2003.
http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.8.8902

The Failure of Poisson Modeling. V. Paxson, S. Floyd. IEEE/ACM Transactions on Networking, 1995.

Traffic Morphing: An Efficient Defense Against Statistical Traffic Analysis Charles V. Wright Scott E. Coull and Fabian Monrose. NDSS, 2009.

Taming the Devil: Techniques for Evaluating Anonymized Network Data S. E. Coull, C. V. Wright, A. D. Keromytis F. Monrose, M. K. Reiter. NDSS, 2008.

In-Network PCA and Anomaly Detection. L Huang, XL Nguyen, M Garofalakis, MI Jordan, A Joseph, N Taft. NIPS, 2007.

A Survey of the State of the Art in Anonymity Metrics Douglas, Kelly, Richard Raines, Michael Grimaila, Barry Mullins and Rusty Baldwin, ACM Workshop on Network Data Anonymization (NDA 2008).

The Devil and Packet Trace Anonymization, Ruoming Pang, Mark Allman, Vern Paxson, Jason Lee. SIGCOMM, 2006.

On Inferring Application Protocol Behaviors in Encrypted Network Traffic, Charles V. Wright, Fabian Monrose, Gerald M. Masson. JMLR, 2006.

Detecting Anomalies in Network Traffic Using Maximum Entropy Estimation, Yu Gu, Andrew McCallum, Don Towsley. Internet Measurement Conference. 2005.

Values

Stephanie Forrest, Steven A. Hofmeyr, Anil Somayaji, Thomas A. Longstaff, "A Sense of Self for Unix Processes," 1996 IEEE Symposium on Security and Privacy, 1996.
http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.47.6145

Henry Hanping Feng, Oleg M. Kolesnikov, Prahlad Fogla, Wenke Lee, and Weibo Gong, “Anomaly Detection Using Call Stack Information”, S&P 2003
http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.13.6179

Values/Volume of content bytes Ke Wang, Gabriela Cretu, Salvatore J. Stolfo "Anomalous Payload-based Worm Detection and Signature Generation" /Proceedings of the Eighth International Symposium on Recent Advances in Intrusion Detection(RAID 2005)
http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.84.9394

Ke Wang, Janak J. Parekh, Salvatore J. Stolfo, “Anagram: A Content Anomaly Detector Resistant to Mimicry Attack”, RAID 2006
http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.86.4769

Janak Parekh, Ke Wang, Salvatore Stolfo, “Privacy-Preserving Payload-Based Correlation for Accurate Malicious Traffic Detection”, LSAD 2006
http://portal.acm.org/citation.cfm?id=1162667

Debin Gao, Michael K. Reiter, and Dawn Song, “Behavioral Distance Measurement Using Hidden Markov Models”, RAID 2006
http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.124.3440

Vertices

Vertices in a graph that represents network connections C. Noble and D. Cook. “Graph-based anomaly detection”. In Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pages 631–636, 2003
http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.101.8966

J. Sun, H. Qu, D. Chakrabarti, and C. Faloutsos, “Relevance search and anomaly detection in bipartite graphs”. SIGKDD Explorations, 7(2), December 2005
http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.125.4449

Vertices in a clique of email users: Wei-Jen Li, Shlomo Hershkop, Salvatore J. Stolfo, "Email Archive Analysis through Graphical Visualization" Proceedings of the 2004 ACM workshop on Visualization and data mining for computer security, 2004.
http://portal.acm.org/citation.cfm?id=1029208.1029229

Hybrids

K.-A. Kim and B. Karp. Autograph: Toward Automated Distributed Worm Signature Detection. In Proceedings of the USENIX Security Symposium, Aug. 2004
http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.94.5342

Kreibich, C., and Crowcroft, J. Honeycomb—Creating Intrusion Detection Signatures Using Honeypots. In Proceedings of the 2nd Workshop on Hot Topics in Networks (HotNets-II) (Nov. 2003 http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.9.6459

C. Kruegel and G. Vigna , “Anomaly Detection of Web-based Attacks”.. 10th ACM Conference on Computer and Communication Security (CCS '03) - #163
http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.4.8030

Singh, S., Estan, C., Varghese, G., and Savage, S. The EarlyBird System for Real-time Detection of Unknown Worms. Tech. Rep.CS2003-0761, UCSD, Aug. 2003
http://www.google.com/url?sa=t&source=web&ct=res&cd=1&url=http%3A%2F%2Fwww.cs.unc.edu%2F~jeffay%2Fcourses%2FnidsS05%2Fsignatures%2Fsavage-earlybird03.pdf&ei=vTA5SojbBtmntgfJxpnaDA&usg=AFQjCNE4heFLlNuct0aCFELUe6eCAwdWvg

Guofei Gu, David Dagon, Xinzhou Qin, Monirul I. Sharif, Wenke Lee, and George F. Riley. “Worm Detection, Early Warning, and Response Based on Local Victim Information”. In Proceedings of the 20th Annual Computer Security Applications Conference (ACSAC 2004), Tucson, Arizona, 2004 http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.95.244

Values/vertices of connection features T. Toth and C. Kruegel, Connection-History Based Anomaly Detection, Proc. IEEE Workshop Information Assurance and Security, June 2002.
http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.20.6226

Value/Volume/Vertices - Characterize meta attacks Robertson, G. Vigna, C. Kruegel, R. Kemmerer, “Using Generalization and Characterization Techniques in the Anomaly-based Detection of Web Attacks”, NDSS, 2006.
http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.109.2599

Bhatkar, S. Chaturvedi, A. Sekar, R., “Dataflow anomaly detection”, S&P 2006
http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.60.1160
23. Shlomo Hershkop, Salvatore J Stolfo, Combining Email Models for False Positive Reduction, KDD 2005. Aug 21-24, Chicago Il.
http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.83.8686

R. Vargiya and P. Chan. Boundary detection in tokenizing network application payload for anomaly detection. In ICDM Workshop on Data Mining for Computer Security(DMSEC), 2003.
http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.114.194&rep=rep1&type=pdf#page=54

PRIVACY ISSUES