Security Event Management: Challenges and Opportunities

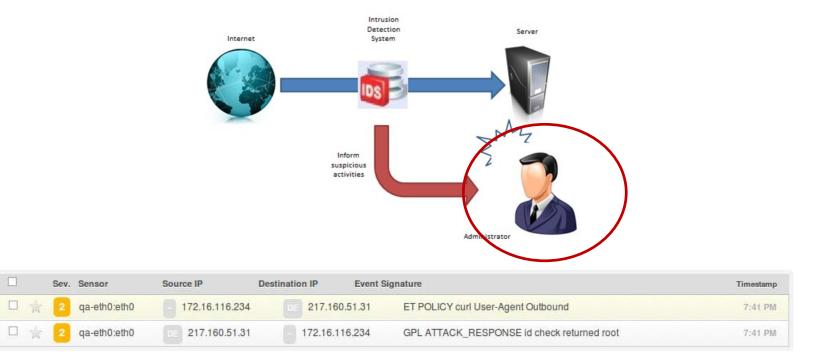
Pratyusa K. Manadhata HP Labs manadhata@hp.com





© Copyright 2014 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice.

Enterprise Security: Point products



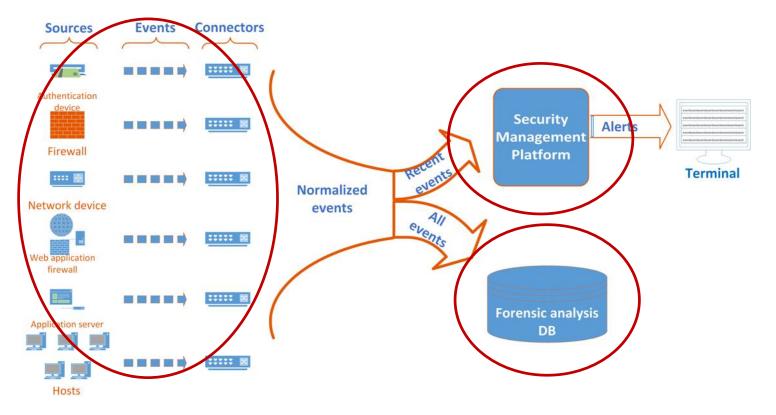


Security information and event management systems (SIEM)

e Preme Constant Constant and Cacle States Cacles States Technicol States (Cacles States Cacles States Cacles States Cacles States Sta						Event Impactor		11							
Mart See: 4 Jan 20125	Andrease United Institutional Applications Indicators Unit Andrease Units and Applications No Mail Andrease Units and Applications No Mail						 A convince Corr Located 	sor eth0:eth0	Source IP Destination IP	Event Signature		Timestamp			
Rater									172.16.116.23			7:41 PM			
Fienager Recotor Time B	tere 8	file Same	Target Law Name	Sepic Attacker Target	Catingery	Page 1									
*Sep 2010 19:00 20 PDF	Top value court little member value to-				Part-Spikater Elecute/Reports Success Menature Security's	_	Event Consult Aven	1	eth0:eth0	217.160.5	16	234 C ATTA	SPOI id k retur	du.	7:41 PM
Also 2010 Photo ALEY	Top value count data montor value co.				Hetilipkatan diwutelikeparan ducasa		8 t	Deribuste 3	9						
					Menature Receiption		Event .	Value .							
4560 202 20 00 00 20 00	Tap value overt data montar value Lu-				Informational Deserts 2		New 1	Query Rutning Time	1						
*Sec 2012 DH 20 20 PDT	Tag value court data montor value co.				Partilephotor EventeReasons Succes		Type 2	Cerelation			• • •				
Ser Write an av	Top value court data montor value to				PartiApplication EvolutioResponse Ducana		Bull Time Application Produced	70H 0H-02-00 C2 (4 C2)			niti	ied U			
			_		Informational Decurity In Practilephonene Medify Configuration Discourse		TemportProtocal			U		CU U			
*Se 20 505 1570	Activitiant entry added	Query Running Tree		> 1525.017.09 (bd%-exec/spl2p-cole)	Annal Applicati		Hutershifty Resource Bytes In								
*Sec 2013 10:00 15 PCF	Additional Success	Query Running Term		+ 16-28-117-29	StatUppkater Molf-Centert Succes Menatoria Japkoter	L	Brites Cult Generator Resource	Contraction Name							
100 203 (200) 15 (21	La contra contra	Cuero Rumeno Time		(10% earl/st/ly.com)	NetAplater Mdh/Cefpater Score	-	Cultural Resource	uters harring rele	ISOT	Courses ID	Destantion ID	Event Cignoture			Transford
and store the second second		Controlled on		(bite een/pi/p.cm)	Nernal Applicate Inschapisature Medif-Content Success	-	Conue D		1301	JULICEIP	Desinialium			rodi	
4 Sep 2012 bit dit 15 PDT	Additionally Success	Query Running Trive		-> 15.25.117.25 (50%-exet.hpl/tp:com)	International Journal		Owner URI		and the second	and the second second				EUL	
* Sep 2010 09 09 10 PDT	Activitial entry added	Query Rurving Time		+ B.B.10.3	Perchastratur PedfyConfiguration (Success Permit Depical		Denan External D Denan Ressure Consul Rens Appropriate C. 1	-eth0:eth0	Source IP False Event Senature For Senature	7:41 PM					
+Sep 2112 19-06-13-927	and the second	Query Running Time		of the second of	Peet/Application (Aud/h/Content /Success										
· 30 217 10 10 10 10	AD D.R. SCHE	Only street into	-	(bith earl hpl/hp.com)	Johanatona Japokaton	-	Consistent Trent Co	1	eth0:eth0	pe 217.160.51.31 172.16 23 TTACK 0 da and pd		7:41 PM			
*Sep.20.0 04:06:15 PDF	Activitial erroy added	Query Running Time		> sh2h2h2028 @dmi-een.tel/second	Peat,Ripkolar MedhyCarlipvator Succes Annal Applicati		Calegory Calegory Spoll-arver								
+Sep 2013 10:00 15 PCF	Additional Success	Query Running Tree		+ 18-28-117-29	PeetApplication (Hadfy/Content /Success	C	Category Behavior Category Technolae								
100 20120-00-05/07	and the second	1000000		(bith early) (by con)	Interview Application Intelligibular Medih/Content Success		Category Device Group								
AND ALL DOM DATE	ASTRUE LICES	Overy Running Time		-> 15-25-107-25 (54% con./s(2p.con)	Anternational Application		Category Outcome Category Object								
454g 300 2406 15 PDF	Activical article attled	Ques Running Time		~ 25.25.217.25 (50%-ear.fpDp.com)	Hatikpkatar HidfyCortpyster Scotte Jame Applat	-	Calepro Tupic Des Threat								
*Sep 200 200 0 0707	Queryliner Query Sutteend	Comectors - Dropping Events	alan i	15 25 128 124 (15 25 116 124 -		-	Hodel Confidence								
		an constant and a constant of	1200	15.25.117.25 (b0%-eas.5pl.7p.las)			Selectly Relevance	10							
*Se 215 84/06 (2 PCT	Querylineer Query Successed	Correctors - Dever - Shari Terre	Max.	10.25 (14.104 (15.25 (14.104) + 15.25 (17.25 (14.104 con hpt.to.com)			Asset Orikulity	1							
🗶 + Say 20.0 (0.05.15 PCF	Query Running Tale	Connectors - Dropping Events				-	Apret								
3.						1	Agent Sevents	1.4	-1						

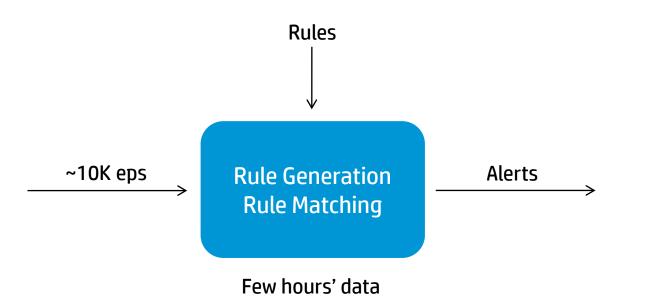


SIEM architecture

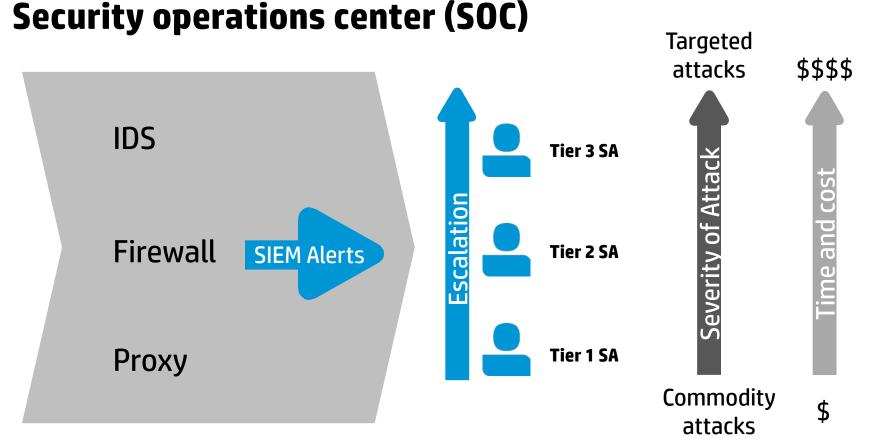




Management platform







6 © Copyright 2014 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice.

(III)

An HP SOC

Hundreds of connector servers A few hundred forensic DBs **Forensic analysis Forensic analysis Forensic analysis** DB DB DB Security Multiple management platforms Management Platform





3 000 000 000 events/day 20 analysts



8 © Copyright 2014 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice.

Operational challenges

Implementing rules – balancing FPs and FNs

Lack of context



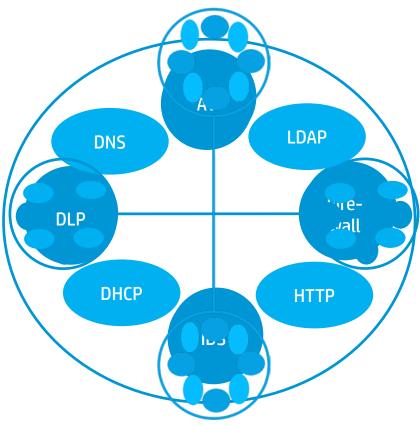
Challenge: Drinking from a firehose

*	Manager Receipt Time 🕇 1	End Time 🗢	Name 🖨	Attacker Address 🖨	Target Address 🗢	Request Url	Device Custom String2	Priority 🗢
	11 Mar 2014 10:23:59 PDT	11 Mar 2014 03:23:58 PDT	Blacklisted DNS Record	16.227.22.197	16.110.135.51	NASSIFG3.americas.hpqcorp.net	Question	5
	11 Mar 2014 10:23:59 PDT	11 Mar 2014 03:23:58 PDT	Blacklisted DNS Record	16.110.135.51	16.225.167.35	kulkapar5.asiapacific.hpqcorp.net	answer	5
	11 Mar 2014 10:23:59 PDT	11 Mar 2014 03:23:58 PDT	Blacklisted DNS Record	16.110.135.51	16.152.82.139	g5w2539.asiapacific.hpqcorp.net	Question	5
	11 Mar 2014 10:23:59 PDT	11 Mar 2014 03:23:58 PDT	DNS Record	16.216.255.16	15.195.192.37	ecowas.org	Question	3

Minutes to decide if an alert or event needs further attention



Challenge: Getting value out of data





Why enterprises collect security data





Research opportunity

Algorithms and systems to identify actionable security information from event data



Research opportunity



Better Enterprise Security



14 © Copyright 2014 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice.

Data-driven security products





AV/IDS/Firewall/..

Signatures/Rules/..



15 © Copyright 2014 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice.

Anti-malware evolution



Static Analysis



Dynamic Analysis



Reputation Analysis

Data analysis: feature extraction/learning/classification/..

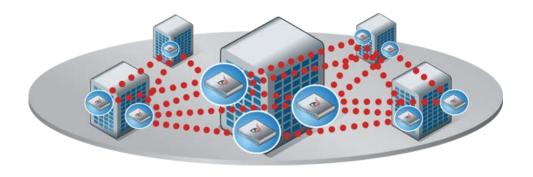


On-Premises analysis

On-site analysis

Fine-grained behavior data

Hardware and virtualization progress enable scaling





Research opportunity

What can we infer from event logs?

How do they compare to on-premises analysis?



Data collection and storage





Scalable analysis

Work with human analysts, not replace them

Things that we took for granted are not true any more



Infer human intent from machine logs

No definition of bad exists – rely on heuristics

Automation is hard

History is not a good indicator of future



Algorithms must learn and evolve

Adversaries adapt

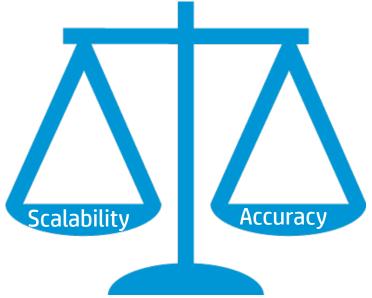
Networks and systems change and fail

People behave unpredictably

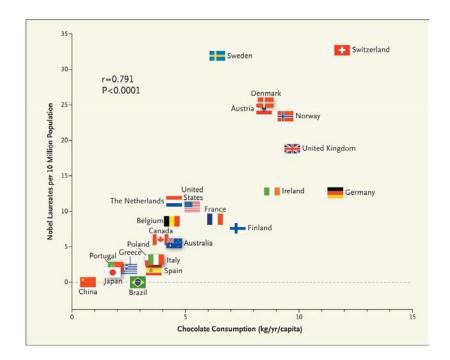


Beware of false positives

Benign events outnumber malicious events



More data = More spurious correlations



Chocolate Consumption, Cognitive Function, and Nobel Laureates, Franz H. Messerli, New England Journal of Medicine, Oct 2012



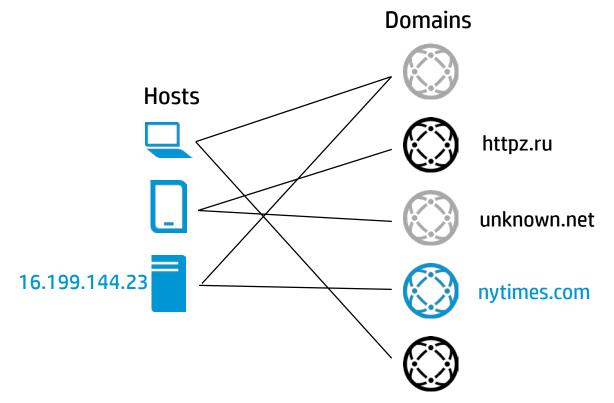


Data minimization vs Serendipity

Privacy-utility trade off

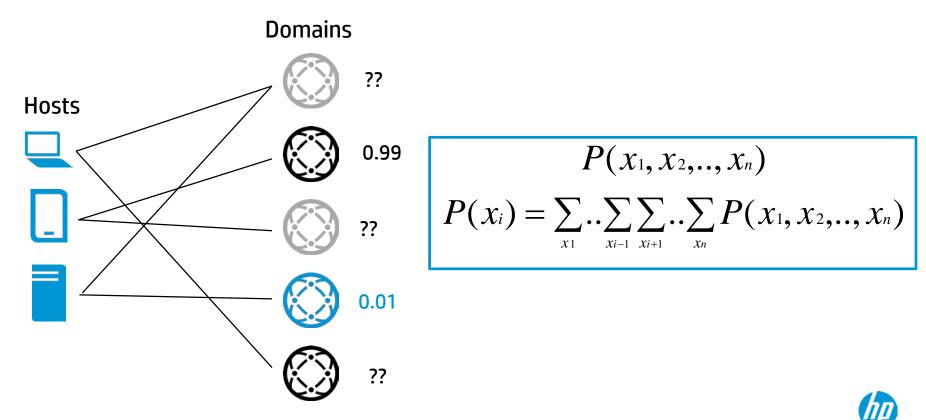


Malicious domain detection





Estimating marginal probability of being malicious



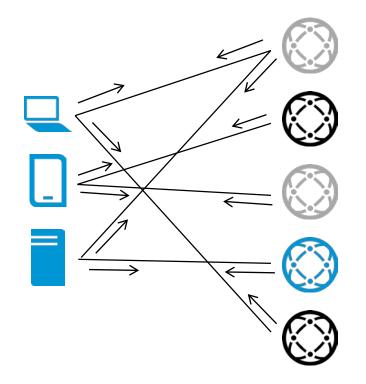
Belief propagation algorithm [P82, YFW01]

Marginal probability estimation in graphs

• NP-complete

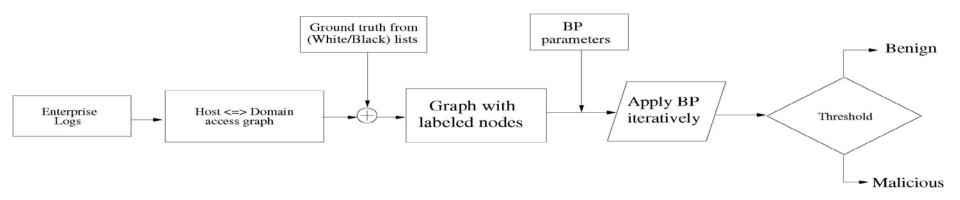
Belief propagation is fast and approximate

• Iterative message passing



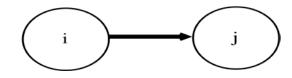


Our approach





Message passing



Message(i \rightarrow j) \propto (prior, edge potential, incoming messages)

$$m_{ij}(x_j) = \sum_{x_i \in S} \phi(x_i) \psi(x_i, x_j) \prod_{k \in N(i) \setminus j} m_{ki}(x_i)$$

$$\uparrow \qquad \uparrow \qquad \uparrow$$
Prior Edge potential Incoming messages



Belief computation

Belief(i) \propto (prior, incoming messages) $b_i(x_i) = K\phi(x_i) \prod m_{ji}(x_i)$ $j \in N(i)$ Normalization constant Prior Incoming messages



HTTP Proxy logs

Logs from a large enterprise

- 98 HTTP proxy servers, 7 months of data
- 1 day's logs : 1.29 billion events
- 2.80M nodes and 27.8M edges

Priors from ground truth (1.45% nodes)

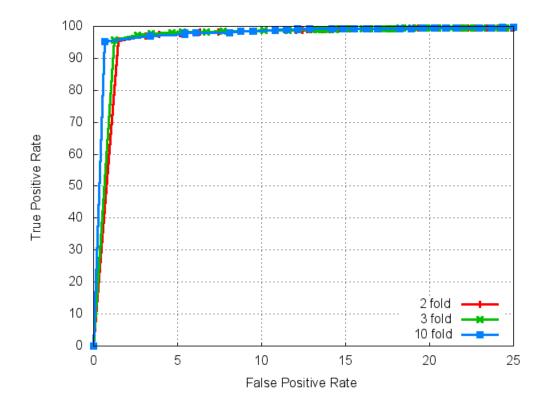
- 21.6K known bad domains: 0.99
- 19.7K known good domains: 0.01
- Unknown hosts and domains: 0.5

Edge potential

	Benign	Malicious
Benign	0.51	0.49
Malicious	0.49	0.51

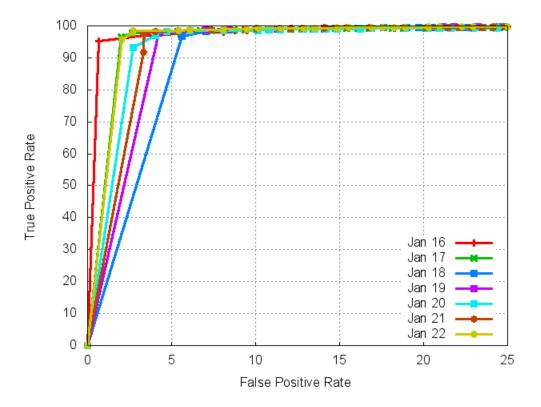


Domain detection ROC plot

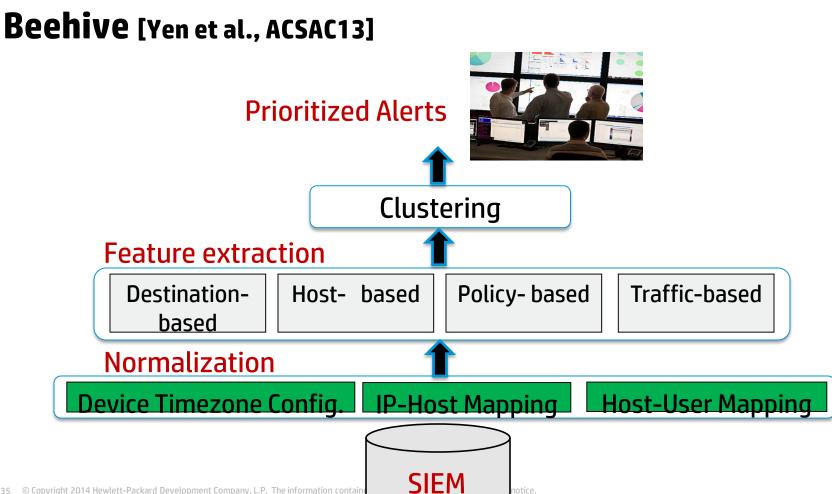




ROC plots for seven days' data







Parting thoughts

A significant Industry problem

Could benefit from academia

Engineering and algorithmic challenges



Thank you



Acknowledgements: Jorge Alzati, Sandeep Bhatt, Stuart Haber, William Horne, Doron Keller, Prasad Rao, and Loai Zomlot

manadhata@hp.com

© Copyright 2014 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice.