SBA

SENTIMENT-BASED BEHAVIOUR ANALYTICS

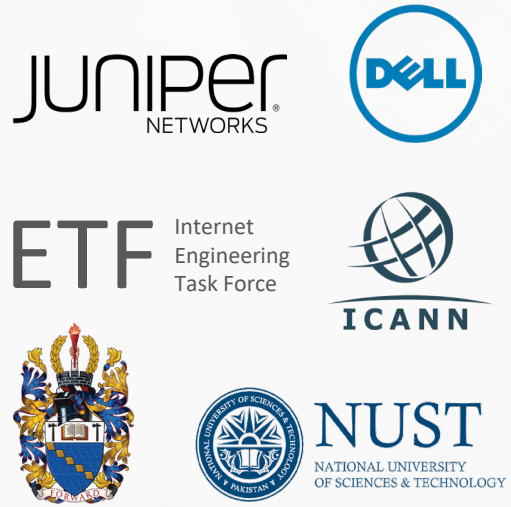
HAFIZ FAROOQ
CYBER SECURITY ARCHITECT / ARAMCO

Upstream Digital Center
Leading Digital Excellence

WHOAM

HAFIZ FAROOQ

- 7 years+ with **Aramco's SOC**
- Splunk Enterprise Architect, ML/AI/LLM Expert
- Previous Companies: Dell, Juniper & other Telcos
- MS (Aston University, UK) | BE (NUST, Pakistan)
- ICANN Fellow | DNS Root Servers, DNSSEC
- IETF Fellow | DNS & DNSSEC Protocols
- Multiple researches in Cyber Security & Routing



Agenda

SENTIMENT-BASED BEHAVIOUR ANALY

- Role of Sentiment Analysis
- Sentiment Datasets for Cyber Security
- Sentiment Analysis Approaches
- Vader Classification Algorithms
- Sentiment-Based Behavior Analytics (SBA)
- Statistics & Results
- Use of Generative AI
- Conclusion

Upstream
Digital Center
Leading Digital Excellence

ROLE OF SENTIME

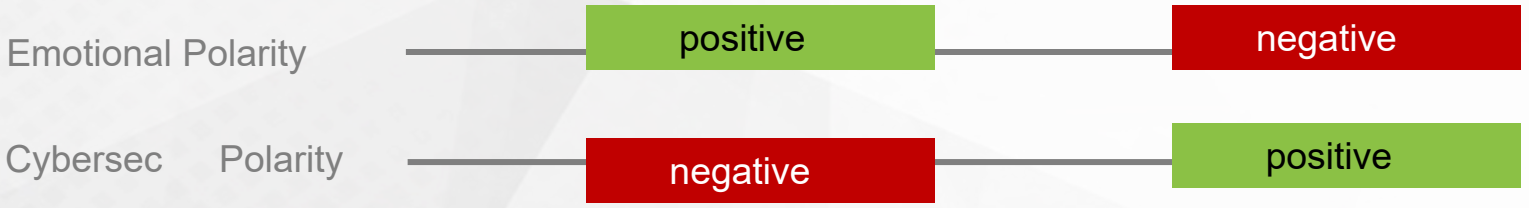
SENTIMENT BASED BEHAVOUR ANALYTICS

Proactive Detection of **negative** cybersecurity sentiments in Saudi Aramco Business Data

EMOTIONS vs SECU

SENTIMENT BASED BEHAVOUR ANALYTICS

Backdoor vs Backbite



Sentiment Lexicon Resources

WordNet, SentiWordNet, SenticNet, MPQA

Cybersecurity Sentiment Normalized Repository | CSNR

Word	Score	Word	Score
Ransomware	-3.9	DarkWeb	-3.7
Hijack	-3.6	Attack	-3.2

1000+ Words


Sentiment Analysis & Insider Threats
Upstream
Digital Center
Leading Digital Excellence

SENTIMENT DATAS

SENTIMENT BASED BEHAVOUR ANALYTICS

Web Browser History | File Names | Process Names

Email Subject/Body | Google Queries

Source Code Variables | DNS Queries

Social Media Access

SENTIMENT ANALYSIS APPRO

SENTIMENT BASED BEHAVOUR ANALYTICS

Approach	Classification	Features	Merits	De-merits
ML Based	- Bayesian Naïve Bayes - Support Vector Machine - DecisionTreeClassifier - Unsupervised Clusters	- Term presence & frequency - Part of speech information - Opinion words and phrase	Create & adapt trained models for specific proposes	Irrelevance to new data as it must be labeled
Lexicon Based	Dictionary Based Corpus based (Statistic or Semantic)	- Dictionary Based - Corpus based Manual	Cover a wider set of words	Fixed sentiment orientation and score of words
Hybrid	Machine Learning & Lexicon Based	- Sentiment lexicon (using public resource) - Sentiment words as features (ML method)	Detection and measurement of sentiment and its less sensitive to change in topic domain	Noisy reviews

Walaa Medhat, Ahmed Hassan, Hoda Korashy Sentiment analysis algorithms and applications: A survey, Ain Shams Engineering Journal, Volume 5, Issue 4, 2014, Pages 1093-1113, ISSN 2090-4479, https://doi.org/10.1016/j.asej.2014.04.011. (https://www.sciencedirect.com/science/article/pii/S2090447914000550)

Lexicon based Analysi

SENTIMENT BASED BEHAVOUR ANALYTICS

VADER Valence Aware Dictionary for sEntiment Reasoning

- VADER relies on a dictionary that maps **lexicons** to valence **scores**
- Valence Scoring **-4** (negative) to **+4** (positive)
- Important Heuristics

  Punctuation | Capitalization | Degree Modifiers | Polarity Shift | Polarity Negation
- Compound Scoring, using Vader formula
- Available in Python

```
# pip install nltk
```

$$x = \frac{x}{\sqrt{x^2 + \alpha}}$$

x = sum of valence scores of constituent words
α = Normalization constant (default value is 15)

Vader Formula

Credits: https://blog.quantinsti.com/vader-sentiment/

SBA Mode

SENTIMENT BASED BEHAVOUR ANALYTICS

Browser History
URL Histories

File Names
Documents & Processes

Data Share
Create / Delete / Update

DNS Datasets
Queries & Answers

Event Collected via
Splunk Universal Forwarders

splunk>

Pre Processing

Machine Learning Toolkit

K-Means Clustering
To find anomalous users

SENTIMENT ANALYSIS PIPELINE

A **Cyber Security Dictionary**
Formulated based on Surveys

NLP

B **Vader Algorithm**
for Sentiment Moderations

C **Sentiment Classification**
Using PCA Dimensionality Reduction

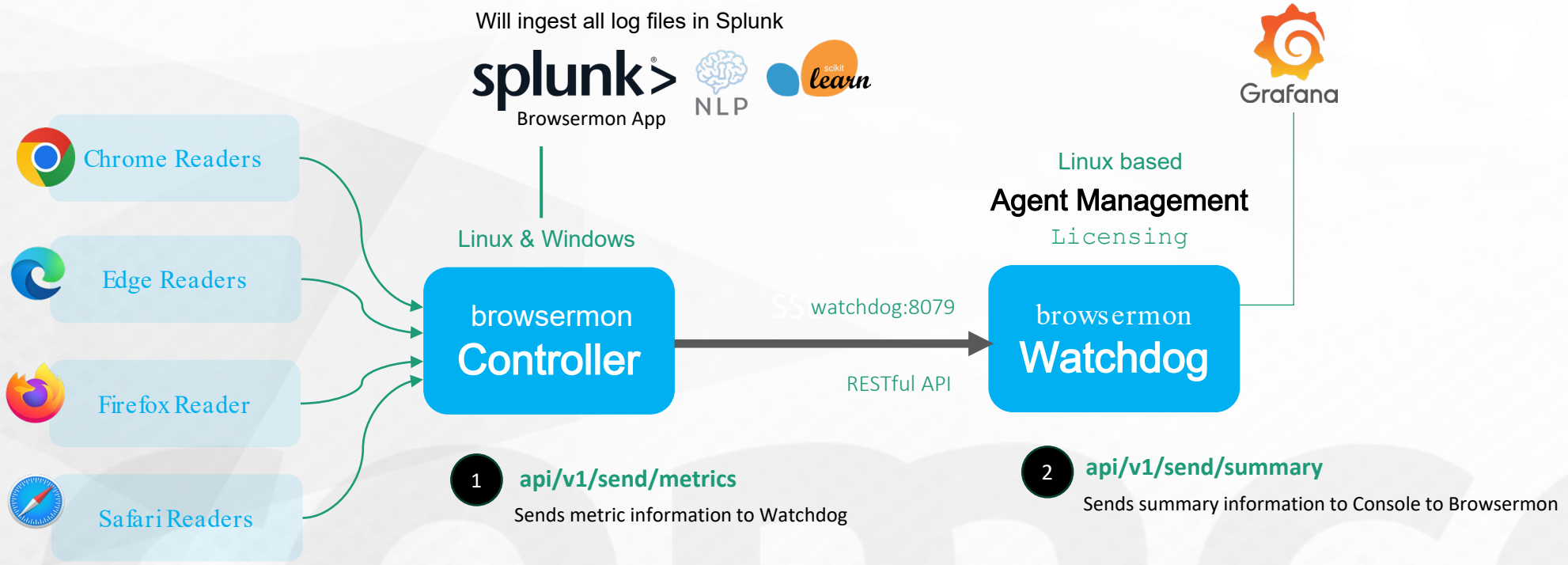
1

Browsing History

BROWSER MON

https://github.com/eunomatix/browsermon
https://browsermon.ai

SENTIMENT SPECIFIC ARTIFACTS



Website URLs | Email Calls | Software Phonehome | Malware C2s

2

File Names

SENTIMENT SPECIFIC ARTIFACTS

Interesting Paths
%USERPROFILE%\AppData
%USERPROFILE%\Desktop
%USERPROFILE%\Downloads

Windows Event ID 4688 — Process Names — cleantextalgorithm

Auditd CVE Events — Binary Names — cleantextalgorithm

SBA

Interesting Paths
/opt/<Applications>
/home/<username>
/root/<username>

3

File Shares

SENTIMENT SPECIFIC ARTIFACTS

Storage Audit Events

Windows Event Viweer

cleantextalgorithm

Network Packet Capture

ZeekFilters

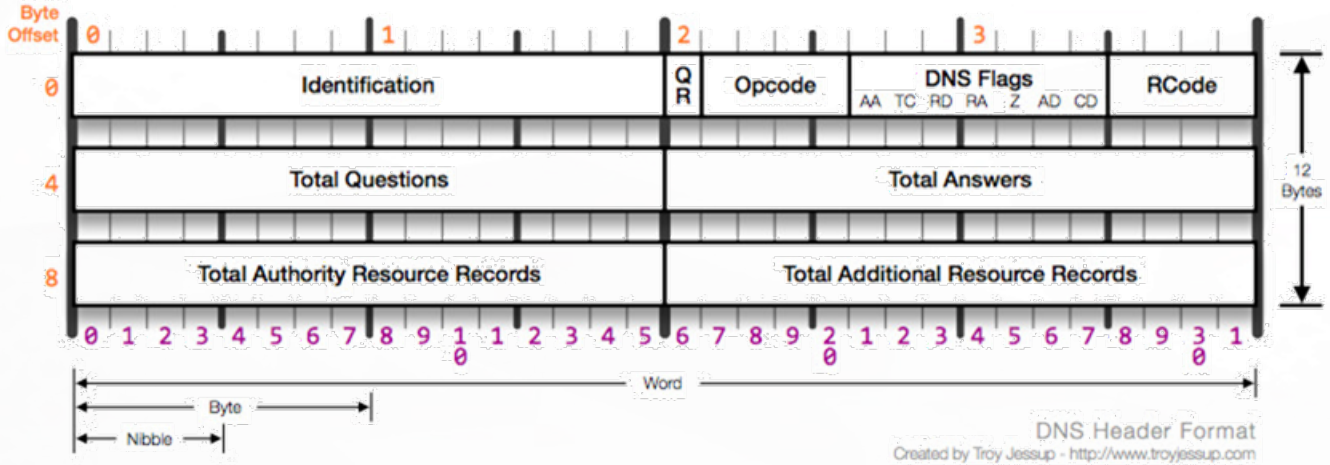
cleantextalgorithm

SBA

4

DNS Queries

SENTIMENT SPECIFIC ARTIFACTS



- Querying Logging of DNS
- DNS Questions & Answers fields
- Extraction of TLDs and Subdomains
- Applying beautifulsoap4 (bs4) for feature extraction
- Applying Sentiment Analysis on the DNS Dataset

Final SBA Mode

SENTIMENT BASED BEHAVOUR ANALYTICS

Browser
URL Histories

File Names
Documents & Processes

Data Share
Create / Delete / Update

DNS Datasets
Queries & Answers

Event Collected via
Splunk Universal Forwarders

splunk>

Pre Processing

Machine Learning Toolkit

Feature Engineering
Extract words in real time

CSNR

A **Cyber Security Dictionary**
Formulated based on Surveys

NLP

B **Vader Algorithm**
for Sentiment Moderations

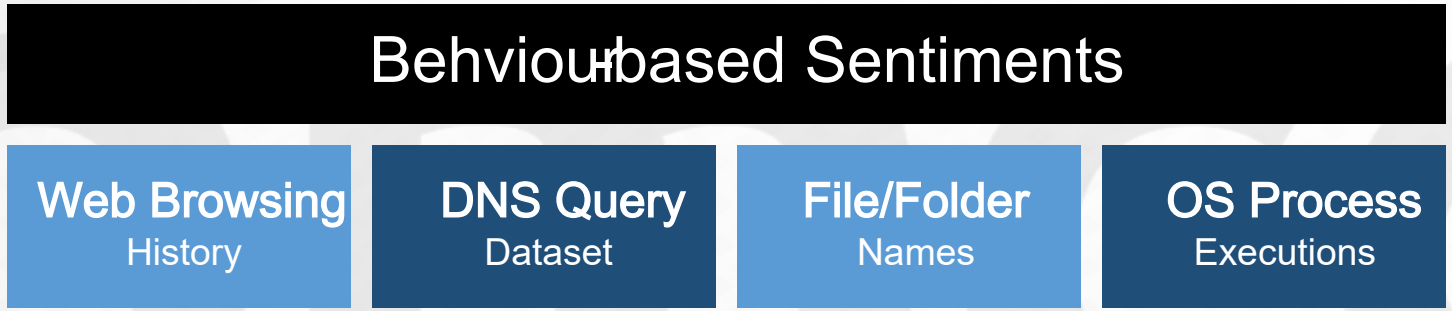
C **Sentiment Classification**
Forwarded to Incident Response Team

SENTIMENT ANALYSIS PIPELINE

OPERATIONAL BENE

SENTIMENT-BASED BEHAVIOUR ANALYTICS

- Early detection of disgruntled employees
- Discovery of Zero-Day Malware, Exploits and Campaigns
- Disclosure of Data Leakage Incidents
- Detection of Compliance violations and new business threats



Sentiments &
Generative AI

Upstream
Digital Center
Leading Digital Excellence

Sentiments &

Generative AI

- ChatGPT Moderation Model for Sentiment Analysis
- Gemini Pro for Sentiment Analysis
- Bard based Sentiment Analysis
- Llama Customized Sentiment Analysis

Punctuation | Capitalization | Degree Modifiers | Polarity Shift | Polarity Negation

Cyber Security
is a key to wards
excellence

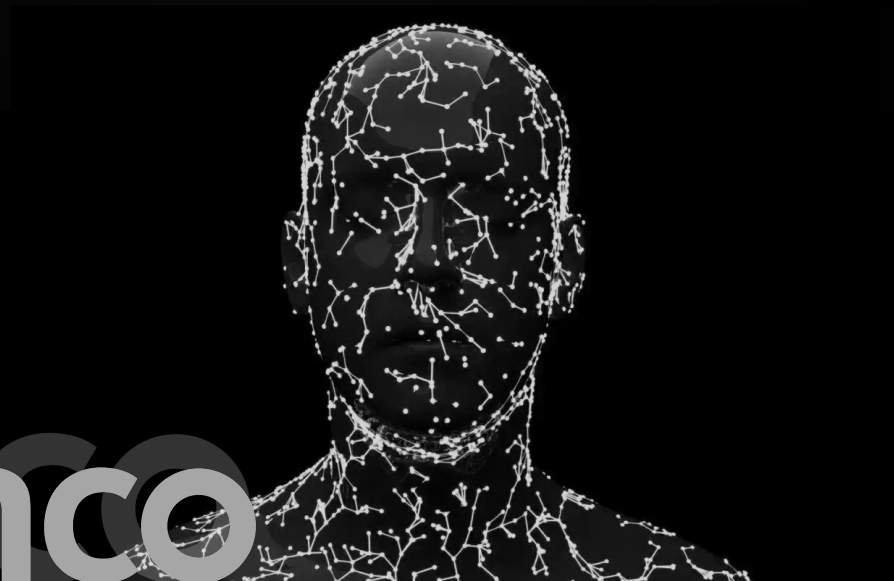
Questions
& Answers

SBA

Sentiment-Based Behavior Analytics

HAFIZ FAROOQ

CYBER SECURITY ARCHITECT / ARAMCO



THANK YOU

SENTIMENT - BASED BEHAVIOUR ANALYTICS

Upstream
Digital Center
Leading Digital Excellence

aramco