GRAPHWAVE

An efficiency improvement for automated vulnerability scans

Tijme Gommers
Northwave B.V.

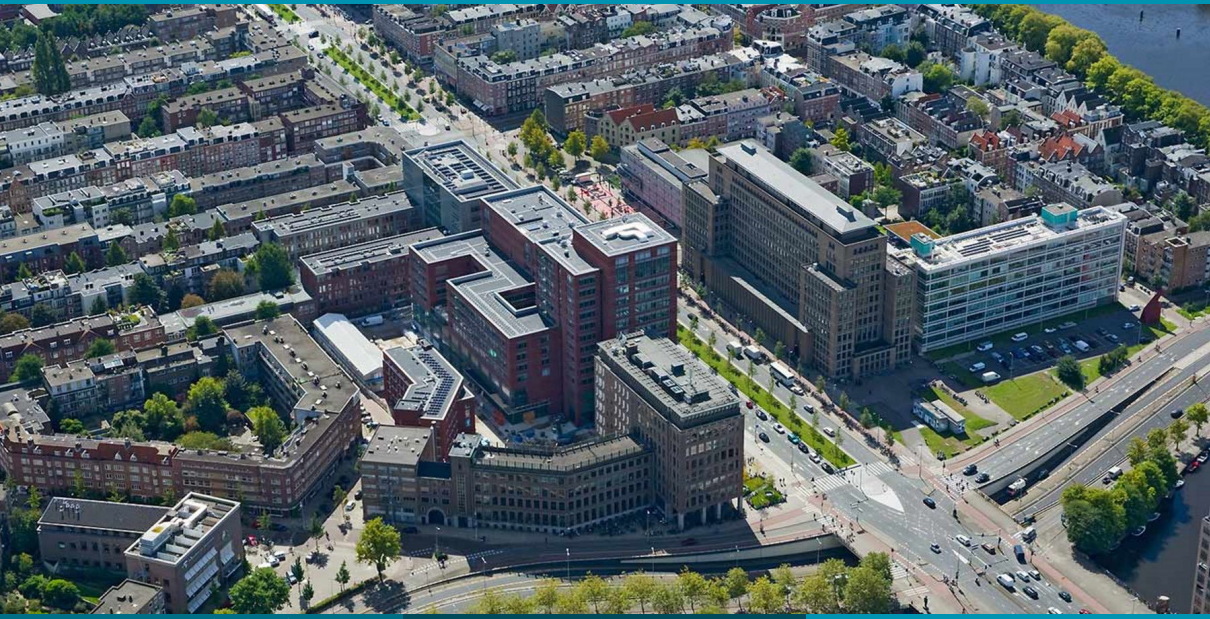
tijme.gommers@northwave.com

2018-06-05

Overview

Introduction
Context

Amsterdam University of Applied Sciences



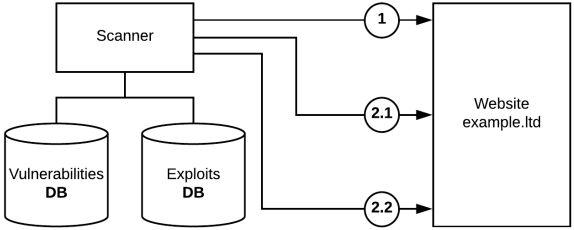
École D'ingénieurs du Monde Numérique

Northwave

Introduction
Research

Automated Web Application Vulnerability Scanners

- 1. Spidering (reconnaissance)
- 2. Crawling (vulnerability scanning)
  - 2.1. Vulnerability database
  - 2.2. Exploit database



Problem

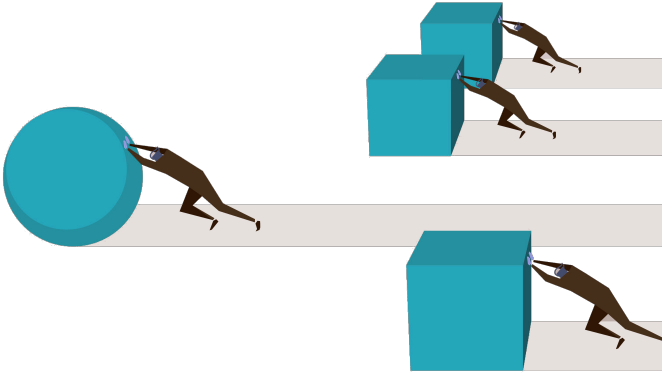
```
Starting: https://nos.nl/artikel/2234370-excelsior-haalt-mahmudov-ook-nieuwe-speler-voor-heracles.html
Starting: https://nos.nl/artikel/2234368-vvv-aast-op-streppel-als-opvolger-van-steijn.html
Starting: https://nos.nl/artikel/2234357-blessure-sterspeler-inspireert-winnende-capitals-in-stanley-cup.html
Starting: https://nos.nl/livestream/npo-nieuws.html
Starting: https://nos.nl/livestream/npo-politiek.html
Starting: https://nos.nl/video/2234358-je-kan-me-bellen-als-je-kanker-krijgt.html
Starting: https://nos.nl/video/2234387-eerste-beelden-neergestort-vliegtuig.html
Starting: https://nos.nl/video/2234356-als-je-niet-dronken-over-straat-mag-lopen-wat-moet-je-dan-in-amsterdam-doen.html
Starting: https://nos.nl/video/2234348-schoolschutter-vs-op-video-jullie-gaan-allemaal-dood.html
Starting: https://nos.nl/video/2234317-belevingsvlucht-toch-wel-erg-dichtbij-en-veel-lawaai.html
Starting: https://nos.nl/video/2234297-de-politiewoordvoerder-vertelt-wat-er-gebeurde-in-schiedam.html
Starting: https://nos.nl/video/2234408-overlast-op-a1-hagelbui-in-midden-van-het-land.html
Starting: https://nos.nl/video/2234344-meer-en-harder-trainen-daar-geloof-ik-niet-in.html
Starting: https://nos.nl/video/2234097-de-schutter-in-[illegible]-agenten.html
Starting: https://nos.nl/video/2234312-babtsjenko-ik-v[illegible]-er-is-gebeurd.html
Starting: https://nos.nl/artikel/2234295-gegijzelde-sch[illegible]-wilde-de-politie-bang-maken.html
Starting: https://nos.nl/artikel/2234[illegible].html
Finished: https://nos.nl/artikel/2234[illegible].html
Starting: https://nos.nl/artikel/2234111-aanslag-op-politie-in-luik-wat-weten-we-van-de-dader.html
Finished: https://nos.nl/artikel/2234369-burgemeester-schiedam-syrische-man-met-bijl-had-psychose.html
Starting: https://nos.nl/artikel/2234156-weinig-retailverdriet-om-blokker-het-is-geen-love-brand.html
Finished: https://nos.nl/artikel/2234409-zidane-vertrekt-bij-real-madrid-er-is-een-andere-coach-nodig.html
Starting: https://nos.nl/artikel/2233925-politieke-powerplay-in-crisis-italie-dit-zijn-de-hoofdrolspelers.html
Finished: https://nos.nl/artikel/2234388-mbo-ers-gaan-voortaan-officieel-studenten-heten.html
Finished: https://nos.nl/artikel/2234341-eikenprocessierups-is-er-vroeg-bij-en-daarom-gevaarlijker-dan-anders.html
Starting: https://nos.nl/nieuwsuur/artikel/2234397-staat-historische-ontmoeting-tussen-kim-jong-un-en-trump-toch-weer-in-de-steigers.htm
Starting: https://nos.nl/video/2234415-een-ritje-door-almere.html
Finished: https://nos.nl/artikel/2234398-dieselverbod-hamburg-gaat-in-zijn-de-dagen-voor-de-dieselauto-geteld.html
Finished: https://nos.nl/artikel/2234336-grote-bedrijven-co2-woon-werkverkeer-en-zakenreizen-halveren-voor-2030.html
Starting: https://nos.nl/video/2234413-stadskantoor-delft-ontruimd.html
Starting: https://nos.nl/video/2234399-moet-je-nog-wel-een-diesel-kopen.html
Starting: https://nos.nl/video/2233769-passend-onderwijs-hoe-werkt-dat.html
Starting: https://nos.nl/video/2229948-zo-werkt-het-kat-en-muisspel-tussen-het-kremlin-en-telegram.html
Finished: https://nos.nl/artikel/2234412-kruis-vanaf-morgen-verplicht-in-beierse-overheidsgebouwen.html
```

36 hours

That is the time an automated scan takes at Northwave

Goal

- Generic open-source solution
- Improve efficiency
- Maintain effectiveness



Efficiency & Effectiveness

- Efficiency = Time to finish
- Effectiveness = Vulnerabilities found

Scope

- Only generic solutions, **not** scanner specific
- Improve efficiency, **not** effectiveness
- Only web application scanners, **not** all scanners
- Only scanners that allow behaviour modifications

Introduction
Questions

Main question

- Which user friendly product can be developed to improve the efficiency of scanners while maintaining effectiveness?

Sub-questions 1

- Which key concept of the scanners is the most promising to improve efficiency?
- Which technology can be used to improve the most promising key concept in an automated way?
- In which user friendly way can the most efficient and effective technology be integrated with scanners?

Sub-questions 2

- Which key concept of the scanners is the most promising to improve efficiency?
- Which technology can be used to improve the most promising key concept in an automated way?
- In which user friendly way can the most efficient and effective technology be integrated with scanners?

Sub-questions 3

- Which key concept of the scanners is the most promising to improve efficiency?
- Which technology can be used to improve the most promising key concept in an automated way?
- In which user friendly way can the most efficient and effective technology be integrated with scanners?

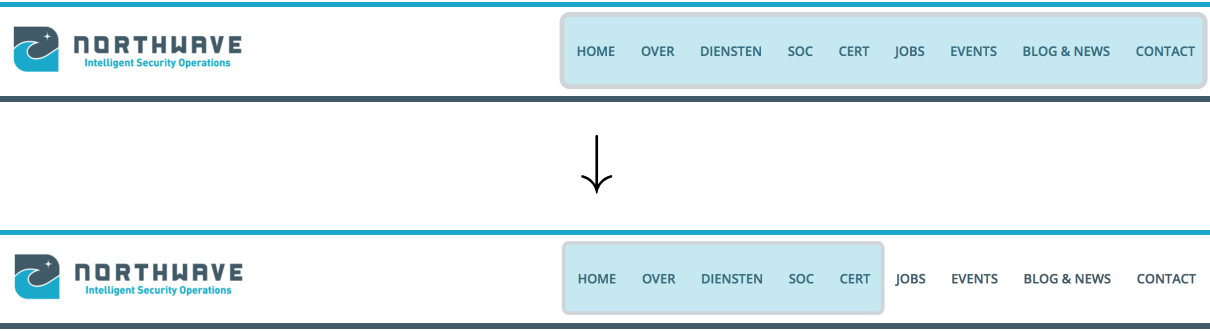
Theory
Automated Web Application Vulnerability Scanners

The three scanners

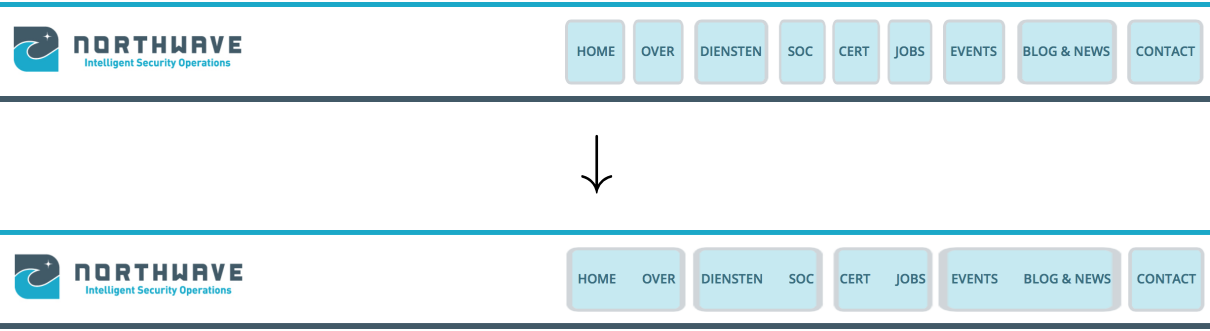
- Burp Suite
- Acunetix
- NYAWC

Theory
Key concepts

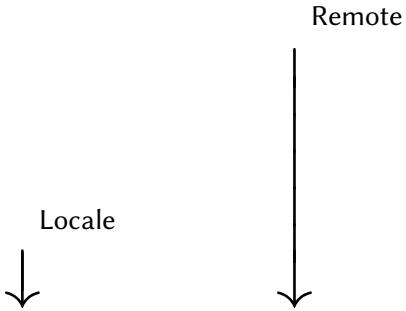
Target scope (reduction)



Multi-threading



Time To First Byte



Persistent HTTP connections

Theory
Measurements

Efficiency

Theorem (efficiency measurement)

The efficiency Y of a scanner is the improved runtime R_i relative to normal runtime R_n.

$$Y(R_n, R_i) = \frac{100}{R_i} \cdot R_n$$

Efficiency example

$$R_n = 36\ hours,\ R_i = 12\ hours$$

$$Y(R_n,\ R_i) = \frac{100}{R_i} \cdot R_n$$

$$Y(R_n,\ R_i) = \frac{100}{12\ hours} \cdot 36\ hours$$

$$Y(R_n,\ R_i) = 300\%$$

Effectiveness

Theorem (effectiveness measurement)

The effectiveness S of a scanner is the amount of vulnerabilities V an improved technology V_i finds on average relative to a normal technology V_n.

$$S(V_n, V_i) = \frac{100}{|V_n|} \cdot |V_i|$$

Effectiveness example

$$V_n = \{vuln1,\ vuln2,\ vuln3,\ vuln4\},\ V_i = \{vuln1,\ vuln2,\ vuln3\}$$

$$S(V_n,\ V_i) = \frac{100}{|V_n|} \cdot |V_i|$$

$$S(V_n,\ V_i) = \frac{100}{|\{vuln1,\ ...,\ vuln4\}|} \cdot |\{vuln1,\ ...,\ vuln3\}|$$

$$S(V_n,\ V_i) = \frac{100}{4} \cdot 3$$

$$S(V_n,\ V_i) = 75\%$$

Promising

Theorem (promising measurement)

The most promising key concept C is a combination of the negative average percentage of the effectiveness S and the average percentage the efficiency Y.

$$C(S, Y) = \frac{200 - (S_l + S_h)}{2} + \frac{Y_l + Y_h}{2}$$

Promising example

$$S_l = 23,\ S_h = 27,\ Y_l = 371,\ Y_h = 464$$

$$C(S, Y) = \frac{200 - (S_l + S_h)}{2} + \frac{Y_l + Y_h}{2}$$

$$C(S, Y) = \frac{200 - (50)}{2} + \frac{835}{2}$$

$$C(S, Y) = 75 + 417.5$$

$$C(S, Y) = 492.5$$

Research
Methodology

Iterative Multi-methodology

- Quantitative as well as qualitative research
- Iterative multi-methodology for each sub-question

Research
Iterations

A promising key concept of the scanners

- Quantitative research
- Three web applications will be tested
- Effectiveness and efficiency should be analysed
- Data should be analysed empirically

A technology to improve the efficiency

- Qualitative and quantitative research
- Find or invent technologies to improve efficiency
- Effectiveness and compatibility should be analysed
- Data should be analysed empirically

A user friendly product

- Quantitative research
- The user-friendliness should be investigated
- Amount of user interactions should be analysed
- Data should be analysed empirically

Results
A promising key concept of the scanners

Analysis

Key concept	Effectiveness	Efficiency
Basis (without modification)	100%	100%
Target scope reduction	83%	138,4%
Multi-threading	100%	16,4%
Persistent HTTP connections	N.A.	N.A.

Table: Promising measurements based on the research design

Analysis

```
Starting: https://nos.nl/artikel/2234370-excelsior-haalt-mahmudov-ook-nieuwe-speler-voor-heracles.html
Starting: https://nos.nl/artikel/2234368-vvv-aast-op-streppel-als-opvolger-van-steijn.html
Starting: https://nos.nl/artikel/2234357-blessure-sterspeler-inspireert-winnende-capitals-in-stanley-cup.html
Starting: https://nos.nl/livestream/npo-nieuws.html
Starting: https://nos.nl/livestream/npo-politiek.html
Starting: https://nos.nl/video/2234358-je-kan-me-bellen-als-je-kanker-krijgt.html
Starting: https://nos.nl/video/2234387-eerste-beelden-neergestort-vliegtuig.html
Starting: https://nos.nl/video/2234356-als-je-niet-dronken-over-straat-mag-lopen-wat-moet-je-dan-in-amsterdam-doen.html
Starting: https://nos.nl/video/2234348-schoolschutter-vs-op-video-jullie-gaan-allemaal-dood.html
Starting: https://nos.nl/video/2234317-belevingsvlucht-toch-wel-erg-dichtbij-en-veel-lawaai.html
Starting: https://nos.nl/video/2234297-de-politiewoordvoerder-vertelt-wat-er-gebeurde-in-schiedam.html
Starting: https://nos.nl/video/2234408-overlast-op-a1-hagelbui-in-midden-van-het-land.html
Starting: https://nos.nl/video/2234344-meer-en-harder-trainen-daar-geloof-ik-niet-in.html
Starting: https://nos.nl/video/2234097-de-schutter-in-                                    -agenten.html
Starting: https://nos.nl/video/2234312-babtsjenko-ik-                                    t-er-is-gebeurd.html
Starting: https://nos.nl/artikel/2234295-gegijzelde-s                                    -wilde-de-politie-bang-maken.html
Starting: https://nos.nl/artikel/2234282-                                    .html
Finished: https://nos.nl/artikel/2234382-                                    .html
Starting: https://nos.nl/artikel/2234111-aanslag-op-politie-in-luik-wat-weten-we-van-de-dader.html
Finished: https://nos.nl/artikel/2234369-burgemeester-schiedam-syrische-man-met-bijl-had-psychose.html
Starting: https://nos.nl/artikel/2234156-weinig-retailverdriet-om-blokker-het-is-geen-love-brand.html
Finished: https://nos.nl/artikel/2234409-zidane-vertrekt-bij-real-madrid-er-is-een-andere-coach-nodig.html
Starting: https://nos.nl/artikel/2233925-politieke-powerplay-in-crisis-italie-dit-zijn-de-hoofdrolspelers.html
Finished: https://nos.nl/artikel/2234388-mbo-ers-gaan-voortaan-officieel-studenten-heten.html
Finished: https://nos.nl/artikel/2234341-eikenprocessierups-is-er-vroeg-bij-en-daarom-gevaarlijker-dan-anders.html
Starting: https://nos.nl/nieuwsuur/artikel/2234397-staat-historische-ontmoeting-tussen-kim-jong-un-en-trump-toch-weer-in-de-steigers.htm
Starting: https://nos.nl/video/2234415-een-ritje-door-almere.html
Finished: https://nos.nl/artikel/2234398-dieselverbod-hamburg-gaat-in-zijn-de-dagen-voor-de-dieselauto-geteld.html
Finished: https://nos.nl/artikel/2234336-grote-bedrijven-co2-woon-werkverkeer-en-zakenreizen-halveren-voor-2030.html
Starting: https://nos.nl/video/2234413-stadskantoor-delft-ontruimd.html
Starting: https://nos.nl/video/2234399-moet-je-nog-wel-een-diesel-kopen.html
Starting: https://nos.nl/video/2233769-passend-onderwijs-hoe-werkt-dat.html
Starting: https://nos.nl/video/2229948-zo-werkt-het-kat-en-muisspel-tussen-het-kremlin-en-telegram.html
Finished: https://nos.nl/artikel/2234412-kruis-vanaf-morgen-verplicht-in-beierse-overheidsgebouwen.html
```

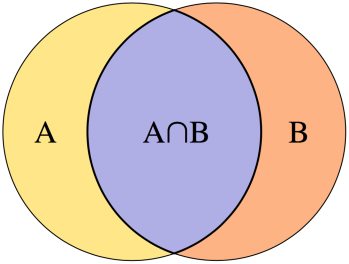
But why?

Why is target scope reduction promising?

Results
A technology to improve the efficiency

HTML tree similarity measure

- Proposed by the Northwave development team
- Jaccard similarity coefficient on HTML trees
- Can be used to measure similarity

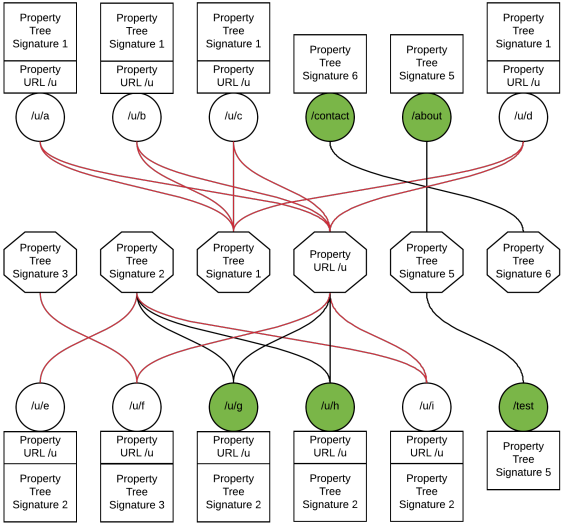


Piecewise response hashing

- Proposed by the Northwave development team
- Piecewise hashing on HTML trees
- Can be used to measure similarity

436:8dslkg48/f48fjoefjwfs:sfkjwfw8w
436:8dslkg48/f48fjoefjwfs:Dfkjwfw8w
Similarity: 99%

A self-developed undirected graph



Analysis

Prototype	Effectiveness	Efficiency
Basis (without modification)	100%	100%
HTML tree similarity measure	65,3%	155%
Piecewise response hashing	75%	148,7%
A self-developed undirected graph	72,3%	156%

Table: Prototype measurements based on the research design

Results
A user friendly product

Analysis

Solution	Extra interactions	Extra seconds
An extension for every scanner	N.A.	N.A.
Crawler that outputs URLs to a file	5	46
Available as an API	3	23
A scanner extension that outputs URLs to a file	4	43

Table: User-friendliness measurements based on the research design

Results
Conclusion

Improving the efficiency of scanners

The best solution is

GraphWave

But a 100% effectiveness is not possible

Product Demo

GraphWave Demo

Future work

Advice for follow-up research

- Similar study to see if GraphWave can be tweaked in such a way that;
  - The graph thresholds and points options are set automatically
  - And the effectiveness can therefore be maintained at a higher level

Questions

Thank you