A Framework for Website Assessment

A. Andreatos
Division of Computer Engineering and Information Science
Hellenic Air Force Academy
Dekeleia Air Force Base, Attica, Greece
a3reatos@hafa.gr, a3reatos@gmail.com

Abstract—The goal of this case study is to assess the websites of institutions similar to ours. The issues of interest are related to site structure, conformance to web design rules, content and performance. We have categorized the criteria and devised methods to measure them; we have examined the websites of interest, assessed the results and drawn conclusions, in order to improve our website look and performance.

I. INTRODUCTION

The web site of an institution, university or company, is its online shop front. Prospect students, customers or partners will interact with it via the Web and make this way their first impression about it. We want to attract our prospect students (or partners or customers) via our website. Two of the most critical factors of a website are: its usability [1] and its performance [2], [3].

Our initial goal was to define some criteria, rate them hierarchically and develop a framework which would allow us to rate the website of the Hellenic Air Force Academy, on the occasion of its reconstruction. Hence, we have grouped five sets of metrics regarding structure, dynamic content, usability, content and performance. Next, we have chosen five European Air Force Academies denoted A1-A5. (Aliases have been used instead of real names.) Then we took measurements, rated the results and finally deduced useful guidelines for the design of our website.

This paper is structured as follows: In section II we present the factors which affect website appearance and performance; in section III we present three mathematical models which we have invented in order to rate some of the aforementioned factors; in section IV we present the five sets of criteria in detail. Section V presents the results and the final rating of the 5 sites under test. In section VI we discuss the results and specify our future work.

II. IMPORTANT PARAMETERS OF WEBSITES

The criteria we considered concern two main issues:

A. Website design rules-guidelines

1. Structure: Size, number of contained objects, number of subpages, update frequency, site map, site search engine etc. (Table I).
2. Dynamic content: Does the page contain frames, scripts, CSS, banners, flashes etc? (see Table II.)
3. Design: Conformance to web design guidelines and conventions (Table III).
4. Content: issues of interest; syllabus, map (how to get there), fees, available languages, admission information, contact information, links (Table IV).

B. Website Performance.

This depends on various factors, the most decisive being [2], [3], [5], [6], [23]:

- Webpage size and complexity [4], [5].
- Server capabilities (such as processor power, memory speed and size, hardware in general and software) [23].
- Networking issues (such as capacity of Internet connections of server and client, Internet state, queues, rush hours, etc). A critical factor is the “smallest pipe” [2], [5] across the server-client path, which works like a bottleneck. We are interested in studying the worst-case scenario here, from the client point of view; hence, we have considered a dial-up client (56 Kbit/sec maximum bandwidth). This is assumed to be the smallest pipe in our case study.
- Operating system, browser, web-server software, caching policies.
- The number of requests accepted by the server.
- Server optimization mechanisms (such as load shedding or load balancing), and network optimization mechanisms (such as Content Delivery Network or edge delivery services) [23].

From the above factors we have assessed webpage size and complexity (in Tables I and II) and response time, Tr (the time it takes to download the webpage to the client’s computer) in Table V. Tr depends on all of the above factors; it is therefore the ultimate performance metric. Another factor directly affecting performance is the HTTP protocol [2], [6], [7], [8], described in RFC 2068 (www.isoc.org). This will not be discussed here and default settings for Windows will be assumed, as is the case with most Internet users.

III. ASSESSMENT MODELS

It is a common tendency to use mathematical models in order to describe physical phenomena. Here we have used three exponential curves, in order to assess some of the parameters under examination. Our first model is the strictly decreasing exponential curve (y = ae−bx). It is frequently used in electronics, for instance, to describe the discharge of a capacitor. Here it is being used to assess the Update Rate [UPD] of a site. The higher the UPD, the better the site. This strictly decreasing behaviour has been rated by the formula: UPD = 10.3exp(-0.027x), shown in Fig. 1. A linear model would give negative results.

The second model is the ALOHA-like exponential curve, which presents a phenomenon with a maximum (optimum) value: y = ax·exp(-x/b). It has been used for

Figure 1. Update rate grading model
assessing parameters which have an optimal value; less is bad, more is also bad. But less is bounded (usually by 0), whereas more is _theoretically_ unbounded. One such parameter is the size of the homepage. The optimal size is often considered to be 40 KB [2], [4]. Values from 20 to 35 and from 45 to 70 KB are considered suboptimal. Too small a homepage will look too poor; it is hard to find an official site below 40KB. On the other hand, a too heavy homepage leads to poor performance: the server cannot process many concurrent requests. Hence, the following formula has been used to rate it: SIZ = 6.8x \exp(-x/40). The corresponding curve is shown in Fig. 2.

The third model is the strictly increasing exponential curve $y=10-10\exp(-0.1x)$. This looks like the capacitor charging curve (voltage versus time). We have used it to rate the number of pages (no.P) of a website; the more the pages, the more the information provided to the interested surfer. A linear model here would lead to excessive values.

Other parameters are binary by nature. For instance: a homepage either provides or it does not provide a site map.

IV. CRITERIA AND RATING METRICS

By “criteria” in this paper we mean the usability or performance parameters, the website characteristics which we examine in order to assess it. Some examples are: whether the site contains broken links or banners, whether it contains outdated information, the number of objects or scripts contained in the homepage etc. By "metrics" we mean the rules we use to assess each parameter, for example: -5 points for each Shockwave Flash movie, -2 points for each broken link, -2 for each Shockwave object, frames, CSS’s and finally render the screen [1], [2], [5]. It is obvious that the more complex the webpage, the slower the site [2], [14]. Hence, the following criteria have been used:

<table>
<thead>
<tr>
<th>#</th>
<th>Criteria</th>
<th>Metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Size (with items)</td>
<td>SIZ=0.68x*\exp(-x/40)</td>
</tr>
<tr>
<td>2</td>
<td>No. pages (mother language website)</td>
<td>no.P=10-10\exp(-0.1x)</td>
</tr>
<tr>
<td>3</td>
<td>Scrolling test, continuous animations</td>
<td>-2 pts each</td>
</tr>
<tr>
<td>4</td>
<td>Update frequency</td>
<td>UPD=10.3\exp(-0.027x)</td>
</tr>
<tr>
<td>5</td>
<td>No. objects</td>
<td>no.O=13\exp(-0.028x)</td>
</tr>
<tr>
<td>6</td>
<td>Site map</td>
<td>YES=+10%</td>
</tr>
<tr>
<td>7</td>
<td>Site search engine</td>
<td>YES=+10%</td>
</tr>
<tr>
<td>8</td>
<td>Broken links?</td>
<td>-2 pts each</td>
</tr>
<tr>
<td>9</td>
<td>Registered in int’l search engines? (in English)</td>
<td>YES=+10%</td>
</tr>
<tr>
<td>10</td>
<td>Own page or subpage?</td>
<td>Own = +10%</td>
</tr>
</tbody>
</table>

B. Dynamic content

The total size of a webpage and the number of contained objects are directly related to its performance. Another very important factor is its complexity, because it affects directly server and client processing times. The server has to process server-side scripts and dynamic webpages written in PHP, ASP.NET or Perl. Similarly, the client has to process client-side scripts written in Java, JavaScript etc., Shockwave objects, frames, CSS’s and finally render the screen [1], [2], [5]. It is obvious that the more complex the webpage, the slower the site [2], [14]. Hence, the following criteria have been used:

<table>
<thead>
<tr>
<th>#</th>
<th>Criteria</th>
<th>Metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Uses frames ?</td>
<td>0 if yes</td>
</tr>
<tr>
<td>2</td>
<td>SCRIPTs</td>
<td>-1 pt each</td>
</tr>
<tr>
<td>3</td>
<td>CSS ?</td>
<td>0 if yes</td>
</tr>
<tr>
<td>4</td>
<td>Shockwave Flash photos</td>
<td>-5 pts each</td>
</tr>
<tr>
<td>5</td>
<td>Shockwave Flash movies</td>
<td>-5 pts each</td>
</tr>
<tr>
<td>6</td>
<td>GIF animations</td>
<td>-2 pts each</td>
</tr>
<tr>
<td>7</td>
<td>Heavy objects</td>
<td>-1/50KB</td>
</tr>
<tr>
<td>8</td>
<td>js</td>
<td>-5 pts each</td>
</tr>
<tr>
<td>9</td>
<td>Banners</td>
<td>-5 pts each</td>
</tr>
<tr>
<td>10</td>
<td>Page is dynamic ? asp, php etc ?</td>
<td>0 if yes</td>
</tr>
</tbody>
</table>

Frames should be avoided if they are not needed [1]. Scripts alone are not bad; lots of scripts though, aggravate site response. CSS’s are bad when they disable people with vision problems [10], [20], and this happens when the specified fonts are less than 12 points (as was the case with the sites examined). Moving images, flashing and continuously moving objects, have an overpowering effect on human peripheral vision and annoy constantly the user [14]. “Web users have learned to stop paying attention to any ads that get in the way of their goal-driven navigation” [10]; it is good to avoid any design or object making the page look like commercial.

C. Web design rules and Usability

Although web design rules and proper authoring style have not been universally adopted, there is some evident progress in web usability [9], which allows us to outline ten of the most important web design rules [1],[4],[9],[10],[11],[12],[13],[14],[15],[16],[17]. Proper writing style is needed when writing for the web. It has been measured that [10],[13],[15],[18],[19]:

![Figure 2. Homepage size grading model](image)

![Figure 3. Number of sub-pages grading model](image)
• 79% of the users scan the page because they find it painful to read too much text on line
• Reading from screen is 25% slower and much more boring than reading from paper
• Web content should have half of the word count of the corresponding paper version.

Internet users are impatient and critical; they want to find what they are looking for as soon as possible. They are not willing to read long pages online. The Web is an informal and immediate medium compared to print; users appreciate an informal, "light" writing style, simple sentences and short paragraphs [4], [15], [18], [19].

Different users have different monitor sizes. Users with wide monitors want to view multiple browser windows simultaneously. Different users have different preferences regarding screen resolution and text size, and they should be left free to use their own preferences; we may not assume that everyone’s screen is 800x600 pixels [11], [19].

Most websites violate usability guidelines, making it difficult for elderly people and users with various problems to use the Internet. The most common problems are related to vision; many sites use tiny fonts or CSS’s which are difficult to read for “senior citizens” (above 65 years-old) [10], [11]. Researches showed that at least 12-point fonts should be used [19], [20], [21]. This rule also applies to links, command buttons, menus etc. Pull-down or pop-up menus, hierarchically walking menus and complex menu structures also make the online lives of elderly and handicapped people hard.

Violation of link conventions is another problem; links which don’t change colour when visited, use of non-standard link colours, links in wrong positions, link overload, linkrot etc. [1], [9], [10], [19], [20]. Table III demonstrates the usability criteria we have considered and their assessment policies.

### TABLE III. Usability Rating

<table>
<thead>
<tr>
<th>#</th>
<th>Criteria</th>
<th>Metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Uses proper content authoring style, Consistency? Avoids over-formatting? Uses short &amp; clear phrases in text, “about us” &amp; window titles?</td>
<td>3, 2, 2, 2 pts respectively</td>
</tr>
<tr>
<td>2</td>
<td>Supports elderly &amp; users with vision problems? Uses limited screen resolution? Enables change of font size?</td>
<td>4, 3, 3 pts respectively</td>
</tr>
<tr>
<td>3</td>
<td>Looks like commercial (ads, banners, animations, pop-ups)?</td>
<td>no: 10, yes: 0, so-so: 5</td>
</tr>
<tr>
<td>4</td>
<td>Avoids too much scrolling?</td>
<td>10 = no scroll, 1 = much scroll</td>
</tr>
<tr>
<td>5</td>
<td>Contains outdated info?</td>
<td>10 if no</td>
</tr>
<tr>
<td>6</td>
<td>Includes “what’s new” or news?</td>
<td>10 if yes</td>
</tr>
<tr>
<td>7</td>
<td>Facilitates navigation? Proper menu design? Easy discovery of info?</td>
<td>3, 3, 4 pts respectively</td>
</tr>
<tr>
<td>8</td>
<td>Respects link conventions? Avoids link overload? Avoids linkrot?</td>
<td>5, 2, 3 pts respectively</td>
</tr>
<tr>
<td>9</td>
<td>Supports low-tech users? Supports alphanumeric browsers? Uses text description of images (ALT)?</td>
<td>10 if yes</td>
</tr>
</tbody>
</table>

D. Content rating

When a prospect student visits a military institution’s website, he/she probably looks for information concerning academic education or military training (“syllabus”), Departments /specialties information, admission information, fees etc. He/she will also appreciate a location map in order to find out how to get there, contact information, addresses and telephone numbers. Additional foreign languages will help international students. Information about academic staff will be useful to colleagues from Universities or other Air Force Academies. Links to similar sites such as other military academies or the EUAFA (European Air Force Academy) and any other useful information (such as lab information, facilities, information about conferences and seminars, visit counter etc.) will also be appreciated [10].

Table IV demonstrates the criteria we used in order to rate the content of each website. Most items are binary. For each additional language we add 3 points, but the maximum is 10. In criteria no. 2, 7 and 10, grade depends on the wealth and accuracy of information.

### TABLE IV. CONTENT RATING

<table>
<thead>
<tr>
<th>#</th>
<th>Criteria</th>
<th>Metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Syllabus</td>
<td>10 if yes</td>
</tr>
<tr>
<td>2</td>
<td>Depts-specialties info.</td>
<td>10 if yes</td>
</tr>
<tr>
<td>3</td>
<td>Academic staff</td>
<td>10 if yes</td>
</tr>
<tr>
<td>4</td>
<td>Academic calendar</td>
<td>10 if yes</td>
</tr>
<tr>
<td>5</td>
<td>Admission info. &amp; fees</td>
<td>5-5 if yes</td>
</tr>
<tr>
<td>6</td>
<td>Additional languages</td>
<td>+3 for each additional lang.</td>
</tr>
<tr>
<td>7</td>
<td>Location info. -map</td>
<td>10 if yes</td>
</tr>
<tr>
<td>8</td>
<td>Contact info.</td>
<td>10 if yes</td>
</tr>
<tr>
<td>9</td>
<td>Links to similar sites</td>
<td>+1 for each link</td>
</tr>
<tr>
<td>10</td>
<td>Other useful info.</td>
<td>Depends on usefulness</td>
</tr>
</tbody>
</table>

E. Performance rating

The performance metric used was the time required for a webpage to appear completely on our browser. We have used a plain dial-up line, first because this case represents the worst case [2], [3], [5], in the sense that the delays are big and the differences get magnified and second because this is the case for most Internet users in Greece (but also worldwide).

In order to rate user satisfaction, we have used the formula: \( SAT = 11 \exp(-0.012T_r) \). This formula follows the “eight second rule” which states that most Internet users abandon a site, if this has not finished downloading within the first 8 seconds [2]. Yet, we have been generous with the slope of the curve, because delays such as 8 sec are unreachable for dial-up users in Greece. Measurements were repeated 10 times on different days and times and results were averaged, in order to compensate accidental errors.

V. THE RESULTS

In order to rate the aforementioned websites, we have gathered the appropriate information, constructed suitable tables and got the following results.

A. Usability criteria

After downloading and examining sites under test, we came out with the results shown in Table VI as far as usability is concerned.

B. Performance

Fig. 5 shows measured response times at 19.2 kbit/s. Inst. A has the worst response time; this is the penalty paid for the wealth of information and multimedia content it provides.
parameters have been used, grouped in four sets. We have

dial-up connection set at 19.2 kbit/s. This modem setting
gEOGRAPHICAL distance from our lab, we have normalized

SAT\(=11\exp(-0.012Tr)\), which converts the
normalized response time into “user

satisfaction”, according to the “8
rule”. Results appear in the

last column of Table V.

### TABLE V. RESPONSE TIME GRADING

<table>
<thead>
<tr>
<th>Institution</th>
<th>Tr (sec) @19.2 kbit/s</th>
<th>Tr (sec) @57.6 kbit/s</th>
<th>SAT(norm) @57.6 kbit/s</th>
<th>SAT @57.6 kbit/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>270</td>
<td>118</td>
<td>119.4</td>
<td>2.6</td>
</tr>
<tr>
<td>A2</td>
<td>200</td>
<td>80</td>
<td>80.084</td>
<td>4.2</td>
</tr>
<tr>
<td>A3</td>
<td>90</td>
<td>46</td>
<td>48.053</td>
<td>6.2</td>
</tr>
<tr>
<td>A4</td>
<td>177</td>
<td>79</td>
<td>77.772</td>
<td>4.3</td>
</tr>
<tr>
<td>A5</td>
<td>140</td>
<td>48</td>
<td>46.591</td>
<td>6.3</td>
</tr>
</tbody>
</table>

Sites A3 and A5 perform relatively well. On the other
hand, A1 performs poorly. Finally, Table VI concentrates
all the results and sums up the final score for each
website.

### TABLE VI. ACCUMULATIVE RESULTS

<table>
<thead>
<tr>
<th>Criteria</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>A4</th>
<th>A5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structure</td>
<td>83</td>
<td>49</td>
<td>55</td>
<td>57</td>
<td>51</td>
</tr>
<tr>
<td>Dynamic content</td>
<td>67</td>
<td>83</td>
<td>76</td>
<td>56</td>
<td>62</td>
</tr>
<tr>
<td>Usability</td>
<td>66</td>
<td>71</td>
<td>61</td>
<td>42</td>
<td>64</td>
</tr>
<tr>
<td>Content</td>
<td>81</td>
<td>52</td>
<td>24</td>
<td>66</td>
<td>54</td>
</tr>
<tr>
<td>SAT (Tr)</td>
<td>3</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>SUM</td>
<td>300</td>
<td>259</td>
<td>222</td>
<td>225</td>
<td>237</td>
</tr>
</tbody>
</table>

Institution A1 scored the best grade (although it had a
heavy page with rich multimedia content), because the
site was carefully designed and full of useful information.
On the other hand, site A3 was rather poor in content and
this is the reason why it scored high in performance. Web
design has a lot of trade-offs, the most important being
that of rich content vs. performance.

### VI. CONCLUSION

This case study rated the websites of five European
Air Force Academies with regard to two kinds of criteria:
usability and performance. The most important usability
parameters have been used, grouped in four sets. We have
measured and rated them using various methods and
models described in the paper.

Through the performed case study a framework for
assessing web sites was developed. Since the criteria used
are commonly accepted and quite general, it is possible

that the framework developed here will be useful for
many similar applications. Nevertheless, some weighting
factors were subjective; these may be adjusted to reflect
various used needs. It is our intention to develop the
proposed framework further, in order to make it more
objective for general use.

This case study allowed us to collect a lot of useful
information and learn useful lessons about website design.

### ACKNOWLEDGMENT

Author wishes to thank Mrs. A. Mykoniati for
formatting the paper.

### REFERENCES


for Optimal Performance”, VP Technology & Operations,

2002.

Greek), Athens, Greece: Davlos, 2000.


performance evaluation system through client observation”, in

http://www.w3.org/Protocols/HTTP/1.0/HTTPperformance.html.

and Lilley C. “Network Performance Effects of HTTP/1.1, CSS1 and PNG”,
http://www.w3.org/Protocols/HTTP/Performance/Pipeline.
June 1997.


[12] Nielsen J., “Top ten guidelines for homepage usability”, May 12,


1997.


and online presentation”,

http://psychology.wichita.edu/optimalsearch.


Design Guidelines Based on Usability Studies with People Age 65

[22] Brownlee N. and Loosely Ch. “Fundamentals of Internet

2002.

Figure 5. Measurement of response time at 19.2 kbit/s

Table V shows the response time, Tr, measured from a
dial-up connection set at 19.2 kbit/s. This modem setting
magnified latency differences.

In order to make the results independent of the
geographical distance from our lab, we have normalized
Tr. Although the figures speak for themselves, we needed
a way to rate the results from 0 to 10. Therefore, we used
the aforementioned formula, 

\[ SAT = 11 \exp(-0.012Tr) \]

which converts the normalized response time into “user
satisfaction”, according to the “8 rule”. Results appear in the
last column of Table V.

![Figure 5. Measurement of response time at 19.2 kbit/s](image-url)