

Trabalho Prático

Análise de Dados em Informática

Engenharia Informática - 3º ano 2º semestre
Ano Letivo 2017/2018

-
- 1. Objetivos**
 - 2. Calendarização**
 - 3. Normas**
 - 3.1 Relatório Final**
 - 3.2 Avaliação**
 - 4. Descrição do Trabalho**
 - 5. Referências Bibliográficas**
-

1. Objetivos

Objetivo Geral:

- Análise de Usabilidade
- Análise ao funcionamento dos servidores VPN do DEI

Objetivos específicos:

- Análise de Usabilidade
 - Definir plano da sessão de avaliação
 - Definir e construir inquérito de satisfação
 - Análise e discussão dos Resultados com recurso ao R
- Análise de fiabilidade
 - Organização do estudo estatístico
 - Análise e discussão dos resultados com recurso ao R
- Escrita do relatório

2. Calendarização

Lançamento das propostas de trabalhos: até 10 de março de 2018

Entrega do trabalho: até 21 de abril de 2018 (23:55)

Defesa e discussão: em data a marcar pelo professor de TP

3. Normas

- O grupo deve ser o mesmo em todos os trabalhos práticos. A defesa e discussão poderão ser realizadas individualmente.
- Deverá ser usado o R como ferramenta de suporte ao estudo estatístico.
- A **data final de ENTREGA** do trabalho é **21 de abril de 2018**, no moodle. Independentemente destes prazos, os grupos deverão ser capazes de, quando o professor o solicitar, reportar o estado de desenvolvimento do trabalho.
- A entrega do trabalho consta de um **relatório e respetivos anexos, inquérito, guia da sessão de avaliação de usabilidade, dados recolhidos, ficheiros de dados** facultados no enunciado e **script do R em ficheiro para as 2 componentes do trabalho (Análise de usabilidade e análise de fiabilidade)**. Deverá submeter todos os documentos num ficheiro compactado. O nome do ficheiro deverá seguir a seguinte notação:

ANADI_YYY_XXX_Nºaluno1_Nºaluno2_ Nºaluno3.zip, onde **YYY** representa a sigla do docente das TP, e **XXX** representa a turma TP.

Exemplo: **ANADI_AIM_3AD_7777777_8888888_9999999.zip**.

- Trabalhos cujo nome não respeite a notação indicada, **serão penalizados em 10%**.
- A **Entrega do trabalho deverá ser submetida no moodle até à data de entrega definida. Não serão aceites trabalhos fora do prazo.**
- A defesa e discussão dos trabalhos decorrerá em dia e hora a marcar por cada professor das teórico-práticas. No dia da apresentação, **TODOS** os elementos do grupo deverão estar presentes. Os elementos ausentes não terão classificação.
- A avaliação do trabalho será realizada por uma equipa de docentes.
- Cada grupo é responsável por gerir o seu processo de desenvolvimento. Dificuldades e problemas deverão ser comunicados atempadamente ao professor das aulas teórico-práticas.

3.1. Relatório

No relatório final deverão ser documentadas todas as fases da Análise de Usabilidade e Fiabilidade realizadas, contextualização do tema, recolha dos dados, organização do estudo estatístico, análise e discussão dos resultados, conclusões e anexos.

3.2. Avaliação

Na avaliação do trabalho serão considerados os seguintes aspetos:

- Contextualização e enquadramento teórico em cada uma das áreas temáticas
- Organização e recolha dos dados
- A análise exploratória de dados com recurso ao R
- A defesa e discussão
- Participação individual de cada um dos elementos

Análise de Usabilidade	40%
• Enquadramento teórico	10%
• Plano da sessão da Avaliação de Usabilidade	5%
• Inquérito	5%
• Análise Exploratória de Dados	20%
 Análise de disponibilidade de servidores VPN do DEI	 40%
• Enquadramento teórico	10%
• Estudo Estatístico	30%
Conclusões	10%
Relatório	10%

Nota: A nota de cada um dos elementos do grupo será definida de acordo com a sua participação. A equipa de avaliação de trabalhos práticos irá validar, no momento da defesa do trabalho, a participação de cada um dos elementos do grupo na concretização dos objetivos do trabalho e do grupo. **Os elementos ausentes não terão classificação.**

4. Descrição do Trabalho

Na realização deste trabalho pretende-se que os alunos desenvolvam atividade no âmbito de duas áreas importantes da Engenharia Informática: a Análise de Usabilidade e a Análise de Fiabilidade.

1. Análise de Usabilidade

Na realização desta componente pretende-se que os alunos desenvolvam o processo de Análise de Usabilidade, com o objetivo de testar e avaliar a usabilidade de sistemas, portais e sites.

Um sistema interativo é constituído por dois elementos principais: a parte computacional, também designada por funcional, e pela parte comunicacional (Hix e Hartson, 1993). A primeira é invisível aos utilizadores e é responsável pelo processamento das ações, pelo esboçar das funções adequadas, e por fornecer ao utilizador o necessário feedback. A segunda, parte comunicacional, é responsável pela comunicação e está a cargo da chamada interface com o utilizador (user interface ou UI), pois é a parte visível. Relativamente ao conceito de usabilidade, este é definido como a eficiência e adequabilidade na

concretização de determinados objetivos por determinados utilizadores (Karray et al., 2008). A definição de usabilidade de um sistema computacional alterou-se ao longo tempo devido à melhor compreensão da interação entre humano e computador e dos fenómenos que rodeiam esse processo de comunicação. Uma das mais conhecidas definições de usabilidade foi apresentada por Nielsen (1993): “a usabilidade apresenta múltiplos componentes e é tradicionalmente associada a cinco atributos: facilidade de aprendizagem, eficiência, facilidade de memorização, reduzida taxa de erros, e satisfação de utilização”. A usabilidade pretende promover todos os aspetos que facilitem a utilização de um sistema computacional. Seguindo os passos definidos por Hix e Hartson (1993) para avaliar a usabilidade, deve ser desenvolvida uma experiência de utilização do sistema a avaliar e organizar sessões de avaliação com participantes onde procedem à recolha de dados. Posteriormente devem realizar a análise dos dados recolhidos e formular as conclusões.

O conceito de *web design* responsivo, proposto por Ethan Marcote (Marcote, 2010), descreve uma técnica na qual o design de um *website* é ajustado automaticamente com base no tamanho dos monitores dos utilizadores. Graças ao design responsivo, os utilizadores podem navegar um site independentemente do dispositivo que usam. O layout e o conteúdo do site serão alterados com base na largura do browser no dispositivo.

No âmbito do processo de desenvolvimento de *websites* as abordagens “*Mobile First Web Designer*” e “*Responsive Web Designer*” surgem como alternativas nestas últimas décadas.

Historicamente, a maioria dos *web designers* e respetivos clientes abordaram o processo de desenvolvimento começando pelo desktop, deixando a componente móvel como um objetivo secundário a ser realizado mais tarde. Mesmo com a ascensão do *design* responsivo, este processo manteve-se. Atualmente, é visível uma tendência crescente para inverter este processo e começar com considerações sobre o desenvolvimento móvel e, em seguida, trabalhar com a versão de desktop maior – ***Mobile First***.

Pretende-se nesta fase testar e avaliar a usabilidade do modelo de interação tendo em consideração o paradigma “*Mobile First*” de um qualquer serviço de *e-commerce*, *e-government* ou *e-business*. Neste âmbito, deverão identificar um *web site* dum serviço existente para diferentes plataformas (desktop, tablet e móvel) de modo a avaliar o impacto do design responsivo na satisfação dos utilizadores e a avaliação do paradigma “*Mobile First*” no panorama nacional.

2. Análise do funcionamento dos servidores VPN do DEI

O DEI mantém vários servidores de rede privada virtual (VPN – Virtual Private Network). Estes serviços permitem aos utilizadores localizados em redes exteriores ao DEI criar uma ligação de rede virtual no seu posto de trabalho equivalente a uma ligação física às redes internas do DEI. Com esta ligação, os serviços das redes do DEI que não estão acessíveis publicamente (externamente) passam a estar disponíveis.

Como acontece com todos os serviços de rede do DEI, é mantido um registo das sessões VPN estabelecidas pelos utilizadores. Com base neste registo de atividades pretendem-se vários tipos de análise ao funcionamento destes serviços e a obtenção de conclusões relevantes para a respetiva administração.

O registo de atividades a ser objeto de análise é fornecido sob a forma de um ficheiro de texto que por motivos de privacidade foi anonimizado (os nomes dos utilizadores e os endereços de rede foram removidos).

Cada sessão é apresentada sob a forma de uma linha de texto, em sequência temporal, contendo os seguintes elementos: servidor, protocolo de VPN, data de início da sessão, hora de início da sessão, hora do fim da sessão, e respetiva duração em minutos.

Extrato exemplificativo:

vsvrv11	PPTP	2016-12-26 17:35	18:33 (58 min.)
vsvrv11	OPENVPN_L2	2016-12-26 17:50	18:36 (46 min.)
vsvrv17	PPTP	2016-12-26 18:29	18:39 (10 min.)
vsvrv17	PPTP	2016-12-26 18:11	18:45 (34 min.)
vsvrv11	PPTP	2016-12-26 18:39	19:00 (21 min.)
vsvrv11	OPENVPN_L2	2016-12-26 19:06	19:14 (8 min.)
vsvrv8	SSTP	2016-12-26 17:57	19:24 (87 min.)
vsvrv17	SSTP	2016-12-26 19:18	19:26 (8 min.)
vsvrv11	PPTP	2016-12-26 19:35	19:36 (1 min.)
vsvrv17	PPTP	2016-12-26 19:24	20:03 (39 min.)
vsvrv8	SSTP	2016-12-26 19:38	20:10 (32 min.)
vsvrv16	SSTP	2016-12-26 16:29	20:12 (223 min.)
vsvrv11	SOFTETHER	2016-12-26 09:00	21:26 (746 min.)
vsvrv17	SSTP	2016-12-26 17:54	21:28 (214 min.)
vsvrv16	PPTP	2016-12-26 15:39	21:28 (349 min.)
vsvrv17	PPTP	2016-12-26 21:24	21:28 (4 min.)
vsvrv16	SSTP	2016-12-26 21:34	21:37 (3 min.)
vsvrv17	PPTP	2016-12-26 16:17	21:37 (320 min.)
vsvrv8	SSTP	2016-12-26 21:40	21:41 (1 min.)
vsvrv8	OPENVPN_L2	2016-12-26 21:39	21:59 (20 min.)

O ficheiro fornecido tem um total de 66383 linhas, abrangendo um período de utilização superior a um ano, de 2016-12-25 até 2018-02-22.

Existem cinco servidores: **vsvrv8; vsvrv10; vsvrv11; vsvrv16; vsvrv17**. Todos os servidores são multiprotocolo, isto é, suportam vários tipos de protocolo de VPN: **PPTP; SSTP; SOFTETHER; OPENVPN_L2; OPENVPN_L3**.

Embora a razão de uma sessão ter uma curta duração possa ter diversas origens, para o efeito deste estudo considera-se que qualquer sessão com duração igual ou inferior a um minuto representa uma falha.

Os dados fornecidos não permitem determinar a duração das falhas, para efeito do cálculo do parâmetro MTBF (Mean Time Between Failures) considera-se que a duração da falha é nula. O MTBF será calculado como o tempo médio entre falhas consecutivas. Para calcular a taxa de falhas num período deve considerar a razão entre o número sessões classificadas como falha e o total de sessões nesse período.

1. Neste estudo devem ser quantificados os seguintes indicadores:

- Número de sessões simultâneas.
- Duração das sessões (excluindo sessões classificadas como falhas).
- Tempo entre falhas consecutivas num mesmo servidor.

Para cada um destes indicadores devem ser calculados o valor mínimo, o valor máximo, o valor médio e o desvio padrão. Adicionalmente, devem ser usadas outras formas de apresentação e organização dos dados. Deve-se também determinar uma estimativa da taxa de falhas.

O estudo destes indicadores deve ser realizado globalmente e parcialmente, nomeadamente:

- Para o conjunto de todos os servidores.
- No conjunto de todos os servidores, para cada tipo de protocolo de VPN.
- Para cada servidor.
- Em cada servidor, para cada tipo de protocolo de VPN.

Cada um dos estudos anteriormente referidos deve ser efetuado para várias janelas temporais:

- Para a totalidade dos dados disponíveis (2016-12-25 até 2018-02-22).
- Para o ano de 2017.
- Para cada mês do ano de 2017.

2. Com base nos dados disponibilizados (ficheiro *vpnsessions.txt*), pretende-se:

a) Determinar a disponibilidade:

- total (durante todo o período de verificação) de cada um dos serviços.
- no ano de 2017 do servidor "**vsrv10**".
- no mês de março 2017 do servidor "**vsrv16**".
- no dia 28 de fevereiro 2017 do servidor "**vsrv17**".

Comente os resultados obtidos.

- b) Determine a função de fiabilidade do servidor "**vsrv8**" nos meses de **janeiro e de fevereiro 2017**. Apresente e analise os respetivos gráficos.
- c) Se λ_i representa a taxa de falhas no dia i do mês de **dezembro de 2017** do servidor "**vsrv16**", em que $i \in \{1, 2, \dots, 31\}$, verifique se a taxa média real de falhas está abaixo dos 0.01 falha/minuto. Opte por um nível de significância que considere mais adequado e justifique a sua decisão.
- d) Valide se existe uma diferença significativa entre os tempos médios entre falhas dos serviços "**vsrv10**" e "**vsrv17**" no período em análise. Use o mesmo nível de significância que usou na questão c) e justifique.
- e) Usando o teste ANOVA, verifique se existe uma diferença significativa entre os tempos médios entre falhas de todos os servidores durante o período total de verificação. Use o mesmo nível de significância que usou na questão c) e justifique.
- f) Retire conclusões com base nos resultados das alíneas anteriores.

Com base no estudo realizado, pretende-se que seja elaborado um conjunto de considerações sobre o desempenho de cada servidor e de cada protocolo de VPN.

Descreva todas as fases necessárias à respectiva concretização. Organize, com base nos resultados obtidos, o estudo estatístico, o mais detalhado e alargado possível. Retire conclusões com base no estudo estatístico realizado. Justifique todas as decisões que tomar.

3. Referências Bibliográficas

KARRAY, F., ALEMZADEH, M., SALEH, J. A., ARAB, M. N., Human-Computer Interaction: Overview on State Art. International Journal on Smart Sensing and Intelligence, 2008.

MARCOTTE, E., Responsive Web Design, <https://alistapart.com/article/responsive-web-design>., May 25 2010.

NIELSEN, J. Usability Engineering. AP Professional, 1993.

HIX, D. e HARTSON, H. R. - Developing User Interfaces: Ensuring Usability Through Product & Process. John Wiley & Sons, inc ,1993.

PIAIRO, E., Desenvolvimento de Interface Inteligente para Suporte à Gestão e Controlo de Produção, Dissertação de Mestrado, FEUP, 2012.

STAPELBERG, R. F. Handbook of Reliability, Availability, Maintainability and Safety in Engineering Design. Springer, 2009.