

Projetos LEI 2020-2021

Fevereiro 2021

1 Sistema de armazenamento baseado em bases de dados chave-valor

Supervisor: João Paulo - jtpaulo@di.uminho.pt e Ricardo Macedo - ricardo.g.macedo@inesctec.pt

Este projeto tem como objetivo desenvolver um sistema de ficheiros, baseado em FUSE, que usa como meio de armazenamento uma base de dados chave-valor (RocksDB). É esperado que se consigam perceber as vantagens e desvantagens em termos de desempenho, escalabilidade e resiliência, de usar esta nova abordagem.

2 Sistema de armazenamento baseado em bases de dados relacionais

Supervisor: João Paulo - jtpaulo@di.uminho.pt

Este projeto tem como objetivo desenvolver um sistema de ficheiros, baseado em FUSE, que usa como meio de armazenamento uma base de dados relacional (PostgreSQL). É esperado que se consigam perceber as vantagens e desvantagens em termos de desempenho, escalabilidade e resiliência, de usar esta nova abordagem.

3 Sistema de armazenamento com compressão transparente de dados

Supervisor: João Paulo - jtpaulo@di.uminho.pt

Este projeto tem como objetivo desenvolver um sistema de ficheiros, baseado em FUSE, que fornece compressão de dados transparente e pode ser utilizado por qualquer aplicação existente.

4 Análise de Aplicações BigData

Supervisor: João Paulo - jtpaulo@di.uminho.pt e Tânia Esteves - tania.c.araujo@inesctec.pt

Este projeto tem como objetivo desenvolver um sistema de tracing, baseado em FUSE, que permita colecionar as operações de armazenamento efetuadas por aplicações de BigData. Estas operações serão guardadas utilizando a ferramenta ElasticSearch de forma a permitir uma análise mais detalhada. Tal análise é fundamental para perceber potenciais problemas de desempenho e funcionamento destas aplicações.

5 Análise de Desempenho de Armazenamento de Kubernetes

Supervisor: João Paulo - jtpaulo@di.uminho.pt

Este projeto tem como objetivo estudar o desempenho de armazenamento oferecido por instalações distribuídas de containers recorrendo à tecnologia Kubernetes. Em maior detalhe, o projeto deverá selecionar diferentes aplicações distribuídas e estudar experimentalmente como estas se comportam ao utilizarem diferentes soluções de armazenamento.

6 Aplicação web para registo histórico de um clube de futebol

Supervisor: jmf@di.uminho.pt

O objetivo deste projeto é o desenvolvimento de uma aplicação para registo de todo o histórico desportivo de um clube de futebol. Trata-se de uma aplicação essencialmente data-centric, em que a organização dos dados se revela crucial para o sucesso do sistema. Pretende-se um registo de todos os jogadores e treinadores (com nome, nome completo, data de nascimento, foto, etc), de todos os jogos oficiais (estádio, adversário, competição, local, data, árbitro, titulares, marcadores de golos, etc). Estes dados todos devem permitir obter diversos tabelas (e.g., tabelas por época, melhores marcadores por competição, jogadores com mais jogos, quadros resumos, etc.). O sistema deverá ainda lançar alertas sempre que algum número especial é atingido por um jogador (e.g., jogo 100, golo 50, recorde de golos numa dada competição). A aplicação deve ser pensada para ser usada pelo gabinete de estatística de um clube ou, de forma mais genérica, pelo gabinete de comunicação. Serve, por exemplo, para fazer sair press releases que recorrem a informação histórica. Deverá ainda ser disponibilizada uma funcionalidade que permita arquivar e classificar fotografias, de modo a que a sua localização possa depois ser facilmente encontrada para efeitos editoriais. Existem já dados para povoar o sistema (cerca de 900 jogadores, 80 treinadores, 3000 jogos) relativos ao SC Braga, mas pretende-se que o sistema seja suficiente genérico para poder ser usado por outros clubes. A equipa de projeto (composta idealmente por 2 ou 3 alunos) deve desenvolver um sistema (web ou desktop) que concretize esta ideia. Espera-se que faça o levantamento dos requisitos (funcionais e não funcionais), a modelação do sistema (diagramas de casos de uso), a conceção e a implementação do sistema e, finalmente, o seu teste. A validação com potenciais utilizadores finais é considerada um aspeto de valorização adicional, embora não seja estritamente obrigatória. Os entregáveis finais são o sistema a funcionar, o código, as bases de dados e a documentação.

7 Using AI for understanding and improving response rates in Patient Reported Outcomes

Supervisor: jmac@di.uminho.pt e hfernandez@ib-s.uminho.pt

With a shift towards a more patient-centered perspective in health care, there is an increase in the use of Patient-Reported Outcome Measurements (PROMs). PROMs are questionnaires that assess health status from patients

perspective and focus on pain, function, quality of life and/or satisfaction. To achieve the goal of evaluating treatments and improving patient care a certain level of response rate is necessary to ensure generalizability and to minimize selection bias of the collected Patient-Reported Outcomes (PROs). The objective is to develop a machine-based learning system to be used to keep in sync with various metrics like the sociodemographic and clinical characteristics of patients/users, time/date for questionnaire response, location of users, bounce rates, exit pages, and more. In this sense, AI has proved to be a powerful platform to conduct different types of tests, and understand the pattern of response rates in different populations. The results of the project can be used for the development of the master's dissertation, which will also have the technical support of the Proef / dstelecom-IB-S chair and the Promptly company.

8 Avaliação de interfaces Linux para I/O

Supervisor: João Paulo - jtpaulo@di.uminho.pt e Ricardo Macedo - ricardo.g.macedo@inesctec.pt

Este projeto tem como objetivo realizar uma avaliação experimental comparativa do desempenho de várias interfaces para I/O disponibilizadas pelo sistema operativo Linux (POSIX, aio, io_uring), sob vários workloads sintéticos e dispositivos de armazenamento (SSDs, memória persistente).

9 Rede 5G virtualizada

Supervisor: alex@di.uminho.pt e hfernandez@ib-s.uminho.pt

A partilha da infraestrutura de rede no caso do 5G permitirá uma considerável redução nos custos de implantação, além de contribuir para solucionar importantes desafios da implantação em áreas urbanas, como potenciais perturbações e a poluição visual da instalação. Essa partilha se baseará em modelos existentes (MORAN e MOCN), mas será complementada com a nova tecnologia de network slicing. Neste sentido, os esforços de padronização e o desenvolvimento de componentes software open-source compatíveis com esses padrões são relevantes para a adoção de modelos de partilha de redes.

Com este projeto pretende-se planejar e iniciar a implementação de uma rede 5G completa, totalmente virtualizada, baseada em componentes de software open-source compatíveis com as especificações da OpenRAN Alliance [1].

Os resultados do projeto poderão ser aproveitados para o desenvolvimento da dissertação de mestrado a qual contará com o apoio técnico dos copromotores da cátedra Proef/dstelecom-IB-S, as empresas Proef Group e dstelecom e o Instituto de Ciência e Inovação para a Bio-Sustentabilidade (IB-S), da Universidade do Minho.

O projeto envolve as seguintes tarefas: 1. Estudo do problema e análise do estado da arte sobre soluções open-source para implementação de redes 5G (RAN e Core) [2]. 2. Definição das diferentes arquiteturas da solução a serem testadas e sua composição: arquitetura de software, interface de ligação entre diferentes componentes, e mapeamento componentes de software e funcionalidade

3GPP. 3. Implementação de uma plataforma MANO e alguns componentes do core 5G. 4. Realização de testes. 5. Escrita do relatório.

Os resultados do projeto poderão ser aproveitados para o desenvolvimento da dissertação de mestrado, a qual também contará com o apoio técnico dos copromotores da cátedra Proef/dstelecom-IB-S, as empresas Proef Group e dstelecom e o Instituto de Ciência e Inovação para a Bio-Sustentabilidade (IB-S), da Universidade do Minho.

[1] ORAN Alliance. Operator Defined Next Generation RAN Architecture and Interfaces. <https://www.o-ran.org/> [2] Netslice Project. Deliverable 3.2 Design and Prototyping of SliceNet Virtualised 5G RAN-Core Infrastructure. <https://bscw.5g-ppp.eu/pub/bscw.cgi/d275928/D3.2-Design>

10 Implementação de uma rede 5G virtualizada baseada no Magma (Facebook Connectivity)

Supervisor: costa@di.uminho.pt e hfernandez@ib-s.uminho.pt

O Magma [1] [2] é uma plataforma de software de código aberto desenvolvida no âmbito da OpenAirInterface Software Alliance [2], pela Facebook Connectivity, uma plataforma lançada pelo Facebook para incrementar o acesso à conectividade a nível global.

Este projeto pretende implementar e testar um ambiente de investigação em redes móveis baseada no Magma, em particular o Magma EPC, e outros módulos de software de código aberto disponíveis. Para tal, será necessário:

1. Estudo do problema e análise do estado da arte sobre soluções open-source para implementação de redes móveis. 2. Definição das diferentes arquiteturas da solução a serem testadas e sua composição: arquitetura de software, interface de ligação entre os diferentes componentes, e mapeamento componentes de software e funcionalidade 3GPP. 3. Implementação da solução baseada no EPC do projeto Magma. 4. Realização de testes. 5. Escrita do relatório.

Os resultados do projeto poderão ser aproveitados para o desenvolvimento da dissertação de mestrado, a qual também contará com o apoio técnico dos copromotores da cátedra Proef/dstelecom-IB-S, as empresas Proef Group e dstelecom e o Instituto de Ciência e Inovação para a Bio-Sustentabilidade (IB-S), da Universidade do Minho.

[1] <https://connectivity.fb.com/magma/> [2] <https://github.com/magma/magma> [3] <https://openairinterface.org/>

11 Ajuste espacio-temporal do Covid-19 Case Fatality Ratio

Supervisor: Raquel Menezes (rmenezes@math.uminho.pt) e Carlos Baquero (cbm@di.uminho.pt)

No decurso da pandemia Covid-19, e em particular em fases de intenso crescimento, torna-se difícil detectar o número de casos apenas com base na testagem por RT-PCR. Uma forma indirecta de inferir o número de casos passa pela

sua extrapolação em função da mortalidade observada e do Case Fatality Ratio (CFR).

Existem, neste momento, boas estimativas para o CFR com base em dados de Wuhan. É, no entanto, sabido que este valor varia de acordo com as regiões e o momento temporal. Factores importantes são a distribuição etária em cada país, e a proporção de gamas etárias infectadas em cada momento. Partindo de uma fórmula já conhecida que calcula o CFR em função da distribuição etária, este trabalho propõe combinar esta com dados de demografia e proporções de infeção ao longo do tempo e em cada país. Pretende-se, assim, automatizar a obtenção de uma série cronológica de valores CFR por país.

Estes resultados podem vir a ser importantes para o projeto internacional, www.coronasurveys.org, onde se apresentam vários indicadores sobre a evolução da pandemia.

12 Uma Nova "Cara" para o Sistema Leonardo

Supervisor: Orlando Belo, obelo@di.uminho.pt

O projeto Leonardo iniciou-se há cerca de dois anos. Pressupunha a criação de um Sistema de avaliação inteligente. Nesta altura, de forma a poder acompanhar as novas funcionalidades e serviços do sistema é necessário conceber, desenvolver e incorporar no sistema uma nova plataforma de interface, capaz de suportar os processos de gestão e de interação dos seus utilizadores no armazenamento, análise e pesquisa de elementos de avaliação e, em particular, nos processos de avaliação de alunos. O objetivo é desenvolver uma interface recorrendo à linguagem VueJS com o auxílio da framework Vuetify. Esta interface terá a função de comunicar com o servidor Web do Leonardo, que está desenvolvido em Python, com a framework Flask, de forma a garantir o acesso e a Gestão de credenciais dos utilizadores do Sistema.

Área de Trabalho: Sistemas de eLearning, Sistemas de Ensino e Avaliação, Plataformas Web, e Interfaces Adaptativos. Ferramentas: VueJS, Vuetify, HTML, JavaScript, Python, Flask e MongoDB.

13 Conversão do Sistema Leonardo para uma Arquitetura Multi-Serviços

Supervisor: Orlando Belo, obelo@di.uminho.pt e António Nestor Ribeiro, anr@di.uminho.pt

O projeto Leonardo tem estado a ser desenvolvido desde há cerca de dois anos para cá no Departamento de Informática. Tem como objetivo fornecer um conjunto de serviços para apoio à avaliação e acompanhamento de estudantes em vários domínios de conhecimento. A estrutura do sistema assenta num conjunto de módulos (aplicações autónomas) coordenadas por um módulo de supervisão centralizado. Na prática, o sistema desenvolvido até ao momento assenta numa estrutura típica de uma aplicação Web monolítica. Neste projeto pretendemos adotar um novo modelo de conceção arquitetural e de desenvolvimento e conceber uma nova versão (para parte) do sistema que fique assente numa plataforma de micro serviços e que possa usufruir das vantagens que este novo modelo de

desenvolvimento nos disponibiliza. Pretende-se intervir na conceção da arquitetura e na identificação dos serviços a disponibilizar de acordo com os padrões arquiteturais mais adequados.

Área de Trabalho: Sistemas de eLearning, Sistemas de Ensino e Avaliação, Micro-Serviços, Plataformas Web . Ferramentas: VueJS, Vuetify, HTML, JavaScript, Python, Flask e MongoDB.

14 Implementação de um Sistema de Interface para uma Document Store de um Sistema de Avaliação de Conhecimento

Supervisor: Orlando Belo, obelo@di.uminho.pt

O sistema Leonardo (este sistema tem estado a ser desenvolvido no Departamento de Informática) tem uma base de dados orientada por documentos (document store) com um conjunto muito diverso de elementos de avaliação, organizado em várias coleções temáticas. Atualmente, todos estes elementos são inseridos na document store através de ferramentas proprietárias associadas com o sistema de gestão de bases de dados que estamos a utilizar. De forma a tornar o sistema independente deste tipo de alimentação e, obviamente, deste tipo de ferramentas, pretende-se que, neste projeto, se desenvolva um sistema de interface para suportar a configuração, definição e inserção dos referidos elementos na document store do sistema, bem como providenciar uma plataforma com serviços de pesquisa de dados que possam ser aplicados sobre a referida base de dados.

Área de Trabalho: Sistemas de eLearning, Sistemas de Ensino e Avaliação, Plataformas Web, Document Stores . Ferramentas: VueJS, Vuetify, HTML, JavaScript, Python, Flask e MongoDB.

15 Implementação de um Sistema de Configuração e Geração Automática de Provas de Avaliação

Supervisor: Orlando Belo, obelo@di.uminho.pt

Neste projeto pretende-se desenvolver um programa que permita definir um conjunto de modelos de provas de avaliação e configurar a forma como cada um dos modelos definidos (e respetivas provas) são apresentadas aos utilizadores de um sistema de acompanhamento e avaliação de estudantes. Posteriormente, o sistema apresentará cada uma das provas definidas de acordo com a sua configuração base, alimentando-as com questões que vai obtendo de um sistema de bases de dados orientado por documentos (uma document store) através do módulo de avaliação (já desenvolvido) do sistema Leonardo este sistema tem estado a ser desenvolvido no Departamento de Informática.

Área de Trabalho: Sistemas de eLearning, Sistemas de Ensino e Avaliação, Plataformas Web . Ferramentas: VueJS, Vuetify, HTML, JavaScript, Python,

Flask e MongoDB.

16 Integração e Expansão de Serviços de Gamificação no Sistema Leonardo

Supervisor: Orlando Belo, obelo@di.uminho.pt

Um grande número de plataformas de gestão tem, hoje, mecanismos e sistemas de gamificação incorporados. Basicamente, tal deve-se, em grande medida, à necessidade de captar a atenção e manter a utilização dos seus serviços pelos seus utilizadores. Dentro do âmbito do projeto do sistema Leonardo (um sistema de avaliação em desenvolvimento do Departamento de Informática) foram concebidos e prototipados vários módulos de serviços de gamificação. Neste projeto, pretende-se desenvolver e integrar, de facto, tais serviços no próprio sistema, e associá-los aos processos de avaliação promovidos pelo sistema, de forma acompanhar o estudante com mensagens de incentivo e dados estatísticos relativos à sua atuação e desempenho.

Área de Trabalho: Sistemas de eLearning, Sistemas de Ensino e Avaliação, Gamificação, Plataformas Web . Ferramentas: VueJS, Vuetify, HTML, JavaScript, Python, Flask e MongoDB.

- 17 Um grande número de plataformas de gestão tem, hoje, mecanismos e sistemas de gamificação incorporados. Basicamente, tal deve-se, em grande medida, à necessidade de captar a atenção e manter a utilização dos seus serviços pelos seus utilizadores. Dentro do âmbito do projeto do sistema Leonardo (um sistema de avaliação em desenvolvimento do Departamento de Informática) foram concebidos e prototipados vários módulos de serviços de gamificação. Neste projeto, pretende-se desenvolver e integrar, de facto, tais serviços no próprio sistema, e associá-los aos processos de avaliação promovidos pelo sistema, de forma acompanhar o estudante com mensagens de incentivo e dados estatísticos relativos à sua atuação e desempenho. Área de Trabalho: Sistemas de eLearning, Sistemas de Ensino e Avaliação, Gamificação, Plataformas Web . Ferramentas: VueJS, Vuetify, HTML, JavaScript, Python, Flask e MongoDB.

Supervisor: Orlando Belo, obelo@di.uminho.pt

A alimentação da document store (base de dados suportada por documentos) do sistema Leonardo (um sistema de avaliação em desenvolvimento do Departamento de Informática), em particular as suas coleções de elementos de avaliação, são alimentadas a partir de um conjunto de dados em formato JSON, que, usualmente, são editados, um a um, através de um editor de texto convencional. Para simplificar o processo de alimentação (especialmente a parte de edição e preparação), precisa-se de implementar um sistema de integração que consiga ler e interpretar os textos (semiestruturados) contendo os diversos tipos de elementos de avaliação com que o sistema Leonardo trabalha. Basicamente, pretende-se construir um programa que leia um texto anotado (de acordo com um conjunto de regras estabelecido), faça a sua interpretação, gere as estruturas JSON correspondentes e, de forma automática, integre essas estruturas na document store do Leonardo. Complementarmente, o programa deverá ser também capaz de registar todas as ações que desenvolver e os erros que detetar.

Área de Trabalho: Processamento de Texto, Plataformas Web, Document Stores. Ferramentas: VueJS, Vuetify, HTML, JavaScript, Python, Flask e MongoDB.

18 Desenvolvimento de um Sistema Web para Banda Desenhada

Supervisor: Orlando Belo, obelo@di.uminho.pt

Neste projeto pretende-se transformar uma plataforma Web existente e adaptá-la para acolhimento e exploração de uma coleção de livros de banda desenhada. A nova plataforma deverá permitir o acolhimento dos diversos elementos que caracterizam um livro de banda desenhada e inseri-los num sistema de bases de dados orientado por grafos. Posteriormente, deverá permitir explorar a base de dados referida, através de formulários de pesquisa e dashboards de gestão específicos.

Área de Trabalho: Plataformas Web, Processamento de Dados, Sistemas de Bases de Dados, Bases de Dados Orientadas por Grafos. Ferramentas: VueJS, Vuetify, HTML, JavaScript, Python, Flask e Neo4J.

19 Desenvolvimento de um Sistema de Integração de Testamentos

Supervisor: Orlando Belo, obelo@di.uminho.pt e Anabela Barros, aldb@ilch.uminho.pt

Neste projeto pretende-se desenvolver uma plataforma Web para acolhimento e exploração de uma coleção de textos de testamentos. A nova plataforma deverá permitir a caracterização e o acolhimento dos referidos textos e inseri-los num sistema de bases de dados orientado por documentos. Posteriormente, deverá permitir explorar a base de dados referida, através de formulários de pesquisa e dashboards de gestão específicos.

Área de Trabalho: Plataformas Web, Processamento de Dados, Sistemas de Bases de Dados, Bases de Dados Orientadas por Documentos. Ferramentas: VueJS, Vuetify, HTML, JavaScript, Python, Flask e MongoDB.

20 Development of Schematic Computer Animated Stories

Supervisor: jmac@di.uminho.pt e nicolas.lori@di.algoritmi.pt

Development of Computer Animated movies is well-known to be extremely expensive, but if the resolution quality is lower (e.g. in videogames) the cost is a lot lower. Following this line of reasoning, it is likely to be possible to make extremely low-quality Computer Animated Stories that will be extremely cheap. In the making of movies and videogames it is a more common trend to search for the highest possible quality of the obtained video, but here, we are looking

in the opposite direction as we are looking for the cheapest approach possible, hence accepting the corresponding lower quality. This option is justified for applications where the goal is to create a large amount of videos, thus requiring a low cost per video. The software tools to be used will be chosen based on the applicant's knowledge, skill, and interests. Software that are likely to be relevant include Blender, Unity, Adobe After Effects, Vicon and Matlab.

21 Sintetização de ações In-vehicle

Supervisor: jmac@di.uminho.pt e tiago.jesus@algoritmi.uminho.pt

Pretende-se com este projeto a geração de dados sintéticos em ambiente In-vehicle, através da ferramenta de animação Blender (não limitado a). O projeto tem como objetivo a criação de vários cenários sintéticos dentro de um automóvel, que irá envolver interação humana. Para a obtenção de modelos humanos, poderá ser utilizado o software Makehuman. O processo de sintetização deverá ser, preferencialmente, automatizado, a fim de gerar cenários sintéticos de forma rápida e personalizada.

22 Sistema de verificação de prescrições de medicamentos

Supervisor: jmac@di.uminho.pt

Pretende-se desenvolver uma aplicação web que dada uma receita médica (medicamento / princípio ativo) apresente informação relevante contida nos folhetos informativos de cada medicamento. Numa segunda fase, será possível verificar se existem interações medicamentosas entre os diferentes medicamentos prescritos.

23 Sistema de monitorização de serviços e servidores

Supervisor: hpeixoto@di.uminho.pt e jamc@di.uminho.pt

Desenvolvimento de um motor de monitorização de aplicações, serviços e hardware. O motor deverá ser capaz de monitorizar diversas métricas de hardware como: ping, diskpace. Monitorizar serviços e aplicações através de uma API, por exemplo. Esta api deverá debitar a informação para uma interface simples e intuitiva que permita de forma simples detetar pontos de falha. Idealmente a interface deverá disponibilizar uma PWA(progressive web app) que permita a instalação em dispositivos móveis, Adicionalmente, devem ser tidas em consideração tarefas de notificação de utilizadores (por exemplo através de email/sms/notificações push). A instalação deste serviço será executada no Centro Hospitalar do Tâmega e Sousa em Penafiel e deverá debitar a informação para um portal externo.

24 Sistema De Suporte à Decisão baseado em Regras Éticas na Saúde

Supervisor: jmac@di.uminho.pt e regina.sousa@algoritmi.uminho.pt

O processo de tomada de decisões, em qualquer que seja a área envolvente, tem sempre por base regras baseadas no consentimento, que foram previamente trabalhadas por vários conselhos. Quando focamos em processos de decisões éticas, chave para o sucesso deste processo está assente na partilha de valores (sejam eles morais ou não) e objetivos. Para além disso, é imprescindível que este processo seja baseado na transparência, na confiança e principalmente na consciência. Todas as argumentações acima descrita tem uma agravante quando associada à saúde de um paciente. Nas instituições de saúde o processo de tomada de decisões, mesmo que muito complexas, deve ser quase imediato e por isso, os profissionais necessitam de suporte tecnológico para que assim o seu dia a dia seja simplificado.

Assim sendo, é necessário um sistema que retire a sobrecarga dos profissionais de saúde, bem como da comissão de ética, no que toca a decisões que por vezes precisam de ser imediatas. Em suma, os problemas associados com a tomada de decisões, nas instituições de saúde, seriam simplificados com um sistema que por base tem a modelação de regras éticas (previamente definidas) e como output tem uma plataforma de suporte à decisão médica. A plataforma de suporte à decisão teria como alicerces processos de Inteligência Artificial, mais especificamente Business Intelligence e Machine Learning.

25 Inteligência Artificial aplicada nas Smart Cities

Supervisor: Luís Miguel Ferreira Rosa, id8123@alunos.uminho.pt e César Analide de Freitas e Silva da Costa Rodrigues, analide@di.uminho.pt

Contextualização

As smart cities são cidades sustentáveis a pensar nas pessoas, onde o progresso social e o bem-estar são o mote para a incorporação de projectos e soluções urbanas. A sustentabilidade da economia, da sociedade e do ambiente dependem desta visão. Por outro lado, a adoção massiva de smartphones, permite transformar uma cidade numa cidade inteligente, contribuindo para atingir um alto nível de desenvolvimento urbano sustentável e melhorar a qualidade de vida dos seus cidadãos. A partir desta tecnologia, suportada por arquiteturas de Internet of Things, podemos coletar dados em tempo real. Por fim, com estes dados é possível aplicar algoritmos de Inteligência Artificial (em inglês, Artificial Intelligence) com o objectivo de entender os padrões dos cidadãos e responder com soluções mais rápidas e de menor custo.

Trabalhos Práticos

Proposta nº1 Rastreamento da mobilidade do indivíduo em período de confinamento

Com este trabalho pretende-se que o grupo de alunos planeie e implemente uma aplicação móvel que, tendo em conta um formulário, permita rastrear a

movilidade dos utilizadores durante o periodo de confinamento.

Proposta nº2 Recomendação de locais seguros durante a pandemia

O objectivo deste trabalho consiste a partir de diferentes fontes de dados, tal como, TomTom, PorData, dados.gov.pt, OpenWeatherMap API construir uma plataforma web que permita recomendar os locais mais seguros, em Portugal, durante a pandemia. Com a construção de gráficos e de mapas interactivos dê um conjunto de recomendações e sugestões ao utilizador.

Proposta nº3 Análise da mobilidade da comunidade global durante a pandemia

O objectivo deste trabalho é explorar, estudar e analisar conjuntos de dados relativo a mobilidade durante a pandemia. Este conjunto de dados fornecidos pela Google (<https://www.google.com/covid19/mobility/>) incluem, entre outros, as tendências de movimento ao longo do tempo por localização geográfica, em várias categorias de locais como retalho e lazer, mercearias e farmácias, parques, estações de transportes públicos, locais de trabalho e residências. Resumindo, os objetivos são (1) explorar, estudar, preparar e analisar os dados a mobilidade; (2) conceber métodos para complementar os conjuntos de dados; e (3) conceber e implementar numa plataforma interactiva.

Ferramentas de Trabalho

Após a escolha do trabalho prático, pretende-se que, o grupo de trabalho realize o seguinte:

Desenvolver uma aplicação móvel em Android (proposta 1) e Web (Proposta 2 e 3) ;

Desenvolver um microserviço com uma API, baseada em REST, usando Spring Framework (<https://spring.io/>);

Implementar o microserviço em ambiente Docker Hub (<https://www.docker.com/>);

Definir e implementar um modelo de dados que permita construir uma base dados relacional (ex. Postgres, MySQL, etc);

Aplicar algoritmos de Machine Learning do tipo Deep Learning (ex. Long Short Term Memory e Convolutional Neural Network) nos dados gerados pela aplicação móvel em linguagem Python usando a biblioteca Keras.

Definir e explicar todas etapas Machine Learning (Pré-processamento de Dados, Treinamento dos modelos e Previsão), incluindo, gráficos, valores que achem necessários na análise dos resultados (<https://www.kdnuggets.com/2018/05/general-approaches-machine-learning-process.html>).

26 Pipeline com capacidade de Extrair, Transformar, Carregar e Apresentar Dados clínicos

Supervisor: jmac@di.uminho.pt e regina.sousa@algoritmi.uminho.pt

As instituições de saúde são organizações complexas vocacionadas para a prestação de cuidados à população. Frequentemente, o princípio da medicina baseada na evidência é comprometido pela falta de evidências sobre mecanismos patogénicos, previsão de risco, falta de recursos e estratégias terapêuticas eficazes. Esse facto é ainda mais evidente em situações de pandemia. Assim sendo, as instituições de saúde estão focadas no investimento de tecnologias

inovadoras que melhorem a experiência não só dos pacientes, mas também dos prestadores de cuidados. O investimento nestas tecnologias fez com que muitos sistemas de armazenamento e processamento de dados surgissem, aumentando exponencialmente os números de dados gerados diariamente. Assim sendo e devido à importância que é dada à gestão e armazenamento de informações do foro pessoal, estes temas são grandes focos de investigação científica. Este projeto tem como principal objetivo ingerir vários tipos de dados (estruturados, semi-estruturados e não estruturados) e submetê-los a processos de extração, transformação e posterior carregamento para uma plataforma onde os mesmos possam ser visualizados. Em suma há dois produtos finais: 1. Pipeline de ETL vocacionado para informações clínicas 2. Plataforma de apresentação de informações que facilitam o processo de tomada de decisões clínicas

27 Gestão de Refsets SNOMED

Supervisor: jduarte@di.uminho.pt e jmac@di.uminho.pt

A Systematized Nomenclature of Medicine Clinical Terms (SNOMED) é uma terminologia clínica internacional multilingue, usada atualmente em mais de 50 países. A Organização portadora da SNOMED disponibiliza toda a estrutura assim como algumas ferramentas para que a SNOMED possa ser instalada e utilizada nas instituições de saúde. Com este projeto, pretende-se criar uma aplicação web que seja capaz de fazer queries à estrutura da SNOMED de forma a que seja possível a gestão (criar, modificar, apagar e consultar) de refsets. Refsets SNOMED podem ser definidos como pequenos grupos de conceitos SNOMED.

28 Desenvolvimento de índices para avaliação da qualidade dos dados

Supervisor: Orlando Belo, obelo@di.uminho.pt e Bruno Oliveira, bmo@estgf.ipp.pt

Analisar dados com baixa qualidade é um problema cada vez mais comum que proporciona a obtenção de conclusões erradas que podem ter efeitos desastrosos para uma organização. A identificação destes problemas, assim como a definição de estratégias para os minimizar, é cada vez mais uma preocupação por parte das organizações. Vários problemas de qualidade de dados podem surgir no contexto de um sistema analítico, como valores ausentes, valores duplicados ou problemas de restrição de integridade. Neste projeto, pretende-se identificar um conjunto de indicadores que permitam determinar a qualidade dos dados utilizados por um sistema analítico. Posteriormente, pretende-se desenvolver um sistema de avaliação da qualidade (web-based) dos dados utilizando os índices estudados.

Área de Trabalho: Sistemas de Data Warehousing, Big Data, ETL, Relatórios/Dashboards. Ferramentas: Microsoft Integration Services, Power BI, Framework .NET/ASP.NET Core

29 Sistema de Recomendação de Exames Médicos para Apoio ao Diagnóstico Clínico Baseado em Sintomas

Supervisor: jmac@di.uminho.pt e diana.ferreira@algoritmi.uminho.pt

O objetivo deste projeto consiste em primeiro lugar na criação de um sistema de recomendação de exames médicos consoante os sintomas inseridos pelo profissional de saúde numa aplicação web ou mobile desenvolvida para efeito. Deste modo, o projeto divide-se em duas vertentes principais: 1) Uso de técnicas de Machine Learning (ML) para recomendar os exames médicos mais adequados para o diagnóstico das condições clínicas associadas aos sintomas manifestados pelo paciente. Este é o principal foco do projeto, onde deve ser feito o benchmarking e o desenvolvimento de modelos de ML. Os dados necessários para o treino do modelo serão fornecidos pelos docentes, mas fica ao encargo dos alunos a preparação desses mesmos dados. 2) Desenvolvimento de uma aplicação Web/Mobile que permita inserir os sintomas para posteriormente alimentarem o sistema de recomendação e devolver a(s) resposta(s) ao utilizador.

30 Detecção de alterações na ocupação do solo Português através de imagens de satélite

Supervisor: António J Esteves, esteves@di.uminho.pt

Utilizando sequências de duas ou mais imagens da mesma localização, obtidas em instantes temporais distintos, pretende-se identificar alterações não sazonais na ocupação do solo Português. Deixa-se à consideração dos alunos optar entre zonas ocupadas por vegetação ou por zonas urbanizadas. As alterações de interesse são aquelas que se revelem permanentes, como por exemplo uma construção nova ou uma zona desflorestada, e não aquelas que sejam variações sazonais (como por exemplo uma zona de floresta em que as árvores de folha caduca perderam as folhas ou a mudança de cor duma zona agrícola). Poderão usar-se imagens/dados de satélite com resolução de 10m (Sentinel-2) ou 3m (Planet).

Referências:

[1] Using Time Series Sentinel-1 Images for Object-Oriented Crop Classification in Google Earth Engine <https://www.mdpi.com/2072-4292/13/4/561/pdf>

[2] Change Detection Based on the Coefficient of Variation in SAR Time-Series of Urban Areas <https://www.mdpi.com/2072-4292/12/13/2089/pdf>

[3] Rapid and EASY Change detection in radar Time-series by Variation coefficient, <https://medium.com/sentinel-hub/reactiv-implementation-for-sentinel-hub-custom-scripts-platform-10aa65fd9c26>

[4] Change Detection for Land Cover Mapping (NASA), <https://appliedsciences.nasa.gov/join-mission/training/english/change-detection-land-cover-mapping>

31 Geração de imagens de alta resolução com Transformers

Supervisor: António J Esteves, esteves@di.uminho.pt

O enorme interesse nas redes neuronais profundas, para resolver problemas de visão por computador e de língua natural, resultou no expoente máximo atual que são os transformers. Os transformer surgiram como uma evolução das redes neuronais recorrentes (RNNs), utilizadas na resolução de problemas de língua natural. As grandes novidades dos transformers são um tamanho (em número de parâmetros) gigantesco, a possibilidade de serem afinados para resolver um de vários problemas (em vez de um único), a capacidade de funcionar em paralelo em vez de sequência (como acontecia nas RNNs), a capacidade de identificar dependências em sequências muito mais longas do que as RNNs e a melhor exploração do mecanismo de atenção. Recentemente começaram a ser aplicados em problemas de visão por computador.

Os transformers foram pensados para identificar dependências de longo alcance em dados sequenciais, nomeadamente em textos longos. Como não se focam em dependências locais, são mais expressivos que as redes convolucionais clássicas, o que os torna interessantes para tarefas que exijam capacidade criativa.

O trabalho proposto neste projeto é assim utilizar um transformer, em conjunto com uma rede neuronal convolucional (CNN), para sintetizar imagens de alta resolução a partir de esboços simples. A CNN aprende um conjunto rico de características das imagens e o transformador combina essas características para criar imagens de alta resolução. O transformer a utilizar pode ser treinado, mas também se pode utilizar um modelo pré-treinado [1].

Referências:

- [1] Taming Transformers for High-Resolution Image Synthesis, <https://arxiv.org/pdf/2012.09841.pdf>
<https://github.com/CompVis/taming-transformers>
- [2] Generating Diverse High-Fidelity Images with VQ-VAE-2, <https://arxiv.org/pdf/1906.00446.pdf>
- [3] Generative Pretraining from Pixels, https://cdn.openai.com/papers/Generative_Pretraining_from_Pixels.pdf
- [4] Hierarchical Autoregressive Image Models with Auxiliary Decoders, <https://arxiv.org/pdf/1903.04933.pdf>
- [5] NLP: BERT and Transformer, <https://jonathan-hui.medium.com/nlp-bert-transformer-7f0ac397f524>
- [6] TensorFlow BERT and Transformer Examples, <https://jonathan-hui.medium.com/tensorflow-bert-transformer-examples-2872e3bbe1e>
- [7] minGPT: A minimal PyTorch re-implementation of the Generative Pre-trained Transformer (GPT) training, <https://github.com/karpathy/minGPT>

32 Detecção de Objetos em Imagens Utilizando Transformers

Supervisor: António J Esteves, esteves@di.uminho.pt

O enorme interesse nas redes neuronais profundas, para resolver problemas de visão por computador e de língua natural, resultou no expoente máximo atual que são os transformers. Os transformer surgiram como uma evolução

das redes neurais recorrentes (RNNs), utilizadas na resolução de problemas de língua natural. As grandes novidades dos transformers são um tamanho (em número de parâmetros) gigantesco, a possibilidade de serem afinados para resolver um de vários problemas (em vez de um único), a capacidade de funcionar em paralelo em vez de sequência (como acontecia nas RNNs), a capacidade de identificar dependências em sequências muito mais longas do que as RNNs e a melhor exploração do mecanismo de atenção. Recentemente começaram a ser aplicados em problemas de visão por computador. É neste contexto que se propõe a utilização e avaliação de um modelo baseado em transformers, o DETR do Facebook, para detetar objetos em imagens (caixa envolvente e tipo de objeto). Poderá optar-se por um modelo pré-treinado ou treiná-lo num conjunto de dados à escolha. A precisão da deteção e o tempo de cálculo/previsão devem ser comparados com os de uma arquitetura mais convencional (por exemplo Faster R-CNN ou Mask R-CNN).

Referências:

- [1] End-to-End Object Detection with Transformers, <https://arxiv.org/pdf/2005.12872.pdf>
- [2] <https://github.com/facebookresearch/detr>
- [3] https://colab.research.google.com/github/facebookresearch/detr/blob/colab/notebooks/detr_demo.ipynb
- [4] <https://paperswithcode.com/method/detr>
- [5] <https://www.youtube.com/watch?v=Vuy7bLq3xbw>
- [6] <https://medium.com/deep-learning-reviews/end-to-end-object-detection-with-transformers-paper-review-e4bc8a803658>
- [7] <https://jonathan-hui.medium.com/tensorflow-bert-transformer-examples-2872e3bbe1e>

33 Segmentação panótica de imagens/vídeo

Supervisor: António J Esteves, esteves@di.uminho.pt

Segmentação de imagens é uma tarefa de visão por computador em que se procura classificar os pixels da imagem numa de várias classes. Existem algumas variantes de segmentação, desde a segmentação semântica (em que apenas se classificam os pixels de acordo com a classe do objeto a que pertencem, mas não se exige delimitar os objetos), passando pela segmentação de instância (em que se pretende identificar que pixels pertencem a cada objeto), até à segmentação panótica (em que se pretende classificar todos os pixels e identificar todas as instâncias de objetos).

O que pretende com este projeto é implementar e testar redes neurais para efetuar segmentação panótica de imagens ou vídeo. A segmentação em vídeo é útil, por exemplo, em sistemas de condução autónoma. As referências incluídas documentam algumas arquiteturas de redes neurais que devem ser analisadas para escolher aquela que vão utilizar na tarefa proposta: VPSNet [1], UPSNet [2], Panoptic-DeepLab [3], DeeperLab [4], SOGNet [5]. Para treinar e testar os modelos recomenda-se utilizar os conjuntos de dados COCO-Panoptic, Cityscapes, Mapillary Vistas, ADE20K ou IDD20K. Para avaliar o desempenho do modelo implementado devem usar-se as métricas Panoptic Quality (PQ) e Parsing Covering (PC).

Referências:

- [1] Video Panoptic Segmentation, <https://arxiv.org/pdf/2006.11339.pdf>, <https://github.com/mcahny/vps>

- [2] UPSNet: A Unified Panoptic Segmentation Network, <https://arxiv.org/pdf/1901.03784v2.pdf>, <https://github.com/uber-research/UPSNet>
- [3] Panoptic-DeepLab: A Simple, Strong, and Fast Baseline for Bottom-Up Panoptic Segmentation <https://arxiv.org/pdf/1911.10194.pdf> <https://github.com/bowenc0221/panoptic-deeplab>
- [4] DeeperLab: Single-Shot Image Parser, <https://arxiv.org/pdf/1902.05093.pdf>
- [5] SOGNet: Scene Overlap Graph Network for Panoptic Segmentation, <https://arxiv.org/pdf/1911.07527.pdf>
- [6] Panoptic Segmentation, <https://arxiv.org/pdf/1801.00868.pdf>
- [7] Awesome Panoptic Segmentation resources, <https://github.com/Angzz/awesome-panoptic-segmentation>

34 Assistente pessoal baseado nas plataformas Notion e Telegram

Supervisor: António J Esteves, esteves@di.uminho.pt

Automação robótica de processos (RPA) é uma técnica em que atividades rotineiras, normalmente executadas por humanos, passam a ser executadas de uma forma automática, simples, e flexível. Neste projeto sugere-se aplicar os princípios de RPA para desenvolver um assistente eletrónico pessoal, tendo como bases a plataforma Notion e o Telegram. Assumindo que existe grande margem de manobra em termos de concretização do trabalho, sugere-se que o Notion seja a plataforma para armazenamento da informação mais persistente do utilizador e o Telegram seja utilizado para suportar a execução do(s) robôs digitais (bots) necessários ao assistente eletrónico. Exemplos de serviços a prestar pelo(s) robô(s): consulta rápida dos emails não lidos, consulta da meteorologia (via OpenWeatherMap), consulta da agenda diária, listagem das notícias do momento (via News API), inserção rápida de conteúdo (texto) no Notion. Para facilitar o desenvolvimento do(s) robô(s) sugere-se a utilização da solução de código reduzido Node-RED.

Referências:

- [1] Bots: An introduction for developers, <https://core.telegram.org/bots>
- [2] Telegram Bot API, <https://github.com/tdlib/telegram-bot-api>
- [3] Node-RED: Low-code programming for event-driven applications, <https://nodered.org>
- [4] Unofficial Python 3 client for Notion, <https://github.com/jamalex/notion-py>
- [5] Automating Notion using notion-py, <https://medium.com/analytics-vidhya/automating-notion-using-notion-py-fa0f2e8f089e>
- [6] Automate Notion with Python, <https://ayushirawat.com/automate-notion-with-python>
- [7] TG2Notion: Telegram to Notion, <https://github.com/5hay/tg2notion>
- [8] Text to notion bot, <https://github.com/dartungar/text-to-notion-bot>
- [9] Notion Markdown Importer, <https://github.com/Cobertos/md2notion>

35 Modelo de Análise Opiniões Híbrido para um Sistema de Ensino Inteligente

Supervisor: Orlando Belo, obelo@di.uminho.pt e Jefferson Magalhães, jeffersonmorais@gmail.com

Um instrumento importante para um processo de ensino-aprendizagem mais personalizado é a análise de opiniões expressas nos feedbacks abertos dos alunos acerca do sistema. Recentemente, no âmbito do projeto Leonardo (projeto interno em desenvolvimento no DI), foi desenvolvido um módulo de análise de opiniões utilizando uma abordagem léxica. Apesar dos resultados satisfatórios do módulo quanto à classificação da polaridade do sentimento (positiva, negativa e neutra) do aluno, a implementação baseada em léxico requer recursos linguísticos poderosos que nem sempre estão disponíveis. Neste sentido, o presente trabalho tem como objetivo incorporar no módulo de análise de opiniões do Leonardo um modelo de aprendizagem híbrido que combine técnicas de machine learning e baseadas em dicionário léxico para prover um processo de classificação da polaridade do sentimento mais assertivo.

36 Análise da eficiência de algoritmos para 'Demosaicing'

Supervisor: bruno.dias@di.uminho.pt

A esmagadora maioria das máquinas fotográficas digitais atuais utiliza apenas um único sensor para recolher os fotões responsáveis pela construção duma imagem digital a cores. Aliás, em máquinas fotográficas digitais "sem espelho"(ou 'Mirrorless') esta é a única opção. O facto de se utilizar apenas um sensor obriga à utilização simultânea de três filtros de cores (vermelho, verde e azul, ou RGB) sobre esse sensor. Cada pixel da imagem é deduzido a partir dum foto-sítio. Existem soluções que aplicam a cada foto-sítio o conjunto dos três filtros (o mais conhecido é o sensor Foveon) e soluções que aplicam apenas um filtro duma cor a cada foto-sítio. Este último método é usado pela grande maioria dos equipamentos atuais. O arranjo que as cores têm no sensor total define um tipo de filtro ou 'Color Array Filter'(CFA). O arranjo mais comum é o Bayer CFA, utilizado por quase todos os fabricantes de máquinas fotográficas digitais com sensores CCD ou CMOS (Nikon, Canon, Sony, Olympus, Pentax, Leica, Panasonic, etc).

Como cada foto-sítio recolhe 'apenas' informação sobre a intensidade duma das três cores, a imagem inicial é formada por pixéis com informação de apenas uma cor. Para que a imagem final, com todos os pixéis com valores RGB, possa ser formada, é preciso estimar o valor da intensidade das cores que não foram filtradas em cada foto-sítio, isto é, num foto-sítio com filtro duma determinada cor é preciso estimar a intensidade das outras duas cores. O processo de construção da imagem final por estimação é conhecido por 'Demosaicing' e é um passo muito importante para a determinação da fidelidade/qualidade das imagens que uma máquina fotográfica digital é capaz de gerar. Ou seja, a aplicação do processo de 'Demosaicing' introduz sempre erros de estimação (alguns des-

ses erros são depois reconhecidos como 'artefactos digitais'). Quanto menores forem os erros maior a qualidade do processo. Existem alguns algoritmos bem conhecidos para fazer 'Demosaicing' e é uma área em constante evolução.

Há alguns anos atrás a Fujifilm introduziu um novo tipo de CFA nos seus sensores X-Trans CMOS. Este CFA utiliza um padrão de cores diferente do Bayer CFA e a empresa sugere que este padrão acarreta vantagens no processo de 'Demosaicing', permitindo a utilização de algoritmos mais eficientes com resultados de maior qualidade.

Assim, o principal objetivo deste trabalho é investigar os mais importantes algoritmos de 'Demosaicing' existentes para CFAs do tipo Bayer e X-Trans, implementá-los e testá-los contra imagens padrão (podem ser utilizadas ferramentas como o Matlab) e fazer uma análise comparativa dos dois tipos de CFA, e dos algoritmos associados, em termos de qualidade (fidelidade) e eficiência (rapidez, consumo de recursos) de 'Demosaicing'. Devem definir-se critérios matemáticos precisos para que a análise comparativa possa ser pertinente. Sempre que possível devem incorporar-se, duma forma crítica, os resultados de projetos ou artigos científicos relevantes nesta área.

Para além do respetivo relatório técnico, deve ser escrito um texto em formato de artigo científico a resumir os resultados do trabalho desenvolvido.

37 Multi-lane Platooning for Traffic-dense Vehicular Highways

Supervisor: bruno.dias@di.uminho.pt

É cada vez mais comum a instalação de dispositivos nos veículos de transporte individual de passageiros, transportes colectivos e de mercadorias, que permite a comunicação entre si e entre veículos e uma infra-estrutura veicular inteligente. Na Europa, as redes veiculares de comunicação (VANET) são normalizadas pelo ETSI e são predominantemente baseadas na tecnologia sem fios IEEE 802.11p. É de esperar que, num futuro próximo, as aplicações VANET passarão a ser suportadas predominantemente sobre redes celulares 5G.

Uma das aplicações VANET sobre a qual mais se tem estudado nas duas últimas décadas é o Platooning. O conceito base desta aplicação tem sido adotado em contextos de comboios de veículos de mercadorias a circular numa única via duma auto-estrada, tendo já sido realizados vários projetos de ID em todo o mundo e que incluem, nalguns casos, implementações experimentais em veículos e estrada reais, ainda que em cenários de experimentação controlados. Na maior parte dos casos, estes projetos estudam e analisam os resultados experimentais obtidos em ambientes simulados, implementados em arquiteturas de simulação específicas para redes veiculares inteligentes/cooperativas.

Nos últimos tempos tem havido uma vontade de aplicar o conceito de platooning em VANET para estradas com mais do que uma via de circulação (Multi-lane Platooning ou Urban Platooning) com tráfego indiferenciado (qualquer tipo de veículo, desde que possuindo equipamento capaz de suportar tecnologias VANET) numa tentativa de ajudar a resolver os problemas graves da elevada densidade de tráfego em auto-estradas de acesso a grande meios urbanos ou hubs de transporte (como os aeroportos). No entanto, é de prever que a

aplicação do conceito de platooning a múltiplas filas de trânsito indiferenciado obriga a adaptações complexas e à redefinição de alguns requisitos funcionais fundamentais bem como à definição de novas limitações à sua aplicação por parte de todos os componentes do sistema (hardware e software).

Em termos genéricos, o principal objetivo deste trabalho será o estudo da aplicabilidade do conceito funcional de platooning a auto-estradas de filas múltiplas de tráfego indiferenciado. Mais detalhadamente, o aluno deve: i) Realizar um estudo detalhado do estado da arte dos projetos de ID atuais e que abordem direta ou indiretamente a temática do Multi-lane Platooning (ou Multi-via); ii) definir os requisitos funcionais, de equipamento e aplicabilidade, dum sistema de platooning multi-via; iii) discutir a especificação comparando criticamente as principais diferenças entre o platooning clássico e o multi-via; e iii) fazer a escrita dum artigo científico que resuma o trabalho desenvolvido e discuta os resultados da análise crítica alcançada.

38 Mecanismos de compressão por padrões para distribuição de áudio

Supervisor: bruno.dias@di.uminho.pt

Os mecanismos de codificação da informação baseada na identificação de padrões de comprimentos crescente têm servido de base à definição de várias normas e desenvolvimento de ferramentas para compressão de dados genéricos sem perda de informação. Esta codificação pode servir para sistemas de armazenamento de dados localmente ou para sistemas de distribuição de dados multimédia.

Historicamente, os desenvolvimentos mais relevantes na compressão de dados genéricos por padrões foram a criação da família de algoritmos LZ (Lempel-Ziv), mais propriamente os algoritmos LZ77, LZ78 e LZW. Esta família de algoritmos gerou, por sua vez, outros algoritmos que otimizam o seu funcionamento num ou mais aspetos técnicos ou que permitem a sua aplicação em simultâneo com outro tipo de tecnologias de codificação da informação.

O principal objetivo deste trabalho é o estudo destes algoritmos e das suas variantes mais relevantes. O estudo deve identificar as diferenças funcionais entre si e os diferentes níveis de desempenho (taxas de compressão, velocidade de execução, complexidade de implementação, etc.) quando aplicados a dados genéricos, a dados de tipos específicos, em contexto de armazenamento local ou para distribuição em rede. O trabalho também deve abordar questões relacionadas com a complexidade da implementação das várias variantes e a sua capacidade para permitirem vários tipos de otimizações. Se possível devem ser construídos pequenos protótipos que exempliquem os vários algoritmos e permitam testar mais facilmente os vários níveis de desempenho.

Do estudo e da análise dos protótipos criados deve resultar a criação dum tutorial.

39 e-tickets em ambiente blockchain

Supervisor: jba@di.uminho.pt e jose.miranda@devisefutures.com

Já a todos nos aconteceu termos um bilhete, voucher, ... que não pudemos utilizar e que também não foi possível vendê-lo a terceiros, devido a tal mecanismo não estar previsto, nem mesmo quando os espetáculos estão esgotados.

Objetivo Definir um e-ticket e utilizar tecnologia blockchain (i.e., DLT Distributed Ledger Technology) para permitir a transação de bilhetes, de acordo com as regras do organizador. Estas regras poderão ser, entre outras: + Os detentores de bilhetes não podem vender o bilhete a terceiros, mas podem devolvê-lo ao organizador do espetáculo, que os aceitará no caso de os conseguir vender; + Os detentores de bilhetes podem vender o bilhete a terceiros, desde que o valor não seja superior ao valor de bilheteira; + Os detentores de bilhetes podem oferecer o bilhete a terceiros, desde que não haja transação financeira.

Bibliografia (inicial) + The Rouge Project Voucher and Note protocol, <https://rouge.network/whitepaper-rouge.pdf> + Projeto em Tecnologia Blockchain, LEI, UMinho, Grupo 72, 2019

40 Assinatura remota empresarial

Supervisor: jba@di.uminho.pt e jose.miranda@devisefutures.com

Em todas as organizações, públicas ou privadas, as soluções de desmaterialização são cada vez mais usadas, começando também a existir pressão para ser adicionado assinatura eletrónica a vários tipos de documentos (PDF, XML, ...) internos da empresa.

Neste projeto pretende-se desenvolver uma plataforma de assinatura remota para documentos internos de uma empresa, constituída por: + Componente de geração de par de chaves e emissão de certificado (por exemplo, utilizando o openssl) para os utilizadores da empresa; + Componente web service de assinatura remota, de acordo com a especificação da Cloud Signature Consortium (<https://cloudsignatureconsortium.org/>).

Para iniciar este projeto, recomenda-se a leitura de: + Especificação CSC em <https://cloudsignatureconsortium.org/resources/download-api-specifications/>

41 Esteganografia

Supervisor: jba@di.uminho.pt e jose.miranda@devisefutures.com

De acordo com https://www.gta.ufrj.br/grad/09_1/versao-final/stegano/index.html "A esteganografia é a arte de esconder informações. Mais especificamente, um meio é escolhido para que uma informação seja camuflada, de modo a evitar que terceiros possam obtê-la. Evidências de seu uso podem ser encontradas ainda na Grécia Antiga. Atualmente, com a popularização dos computadores e da Internet, a chamada esteganografia digital vem ganhando espaço, utilizando imagens, arquivos de áudio e a própria Internet como meios esteganográficos."

Neste projeto pretende-se utilizar técnicas de esteganografia para a criação de uma solução anti-falsificação de baixo custo que possa ser aplicada a produtos do dia a dia (por exemplo, uma garrafa de vinho). Esta solução deve ainda permitir validar online, através de uma app, se o produto é verdadeiro ou falso.

Para iniciar este projeto, recomenda-se a leitura de: + Relatório do projeto de Esteganografia, Eng. Segurança, UMinho, 2019; + Avaliação de algum trabalho já feito na área, como por exemplo <https://www.systechone.com/product/real-time-product-authentication-unisecure/>

42 Sistema de Certificação de Atributos Profissionais - integração

Supervisor: jba@di.uminho.pt e jose.miranda@devisefutures.com

Contexto Portugal, por intermédio da AMA (Agência para a Modernização Administrativa), tem desenvolvido um conjunto de projetos inovadores na área da desmaterialização de documentos e desburocratização de serviços, que se encontram na vanguarda do que é feito a nível europeu e mundial. Um desses projetos, o SCAP, está em fase de credenciação de acordo com o regulamento UE 910/2014 (regulamento eIDAS) e permite: + Assinatura eletrónica qualificada na qualidade, através da adição de atributos empresariais a um documento, reconhecendo a qualidade de quem assina.

Objectivo Com estes projetos pretende-se utilizar as APIs disponíveis para comunicar com o SCAP, para desenvolver add-ins/módulos/plug-ins para vários sistemas, permitindo que as funcionalidades do SCAP sejam mais facilmente acedidas nas ferramentas utilizadas pelos cidadãos no dia a dia.

Projetos As ferramentas nas quais se pretende integrar as funcionalidades do SCAP, através de add-ins/módulos/plug-ins, são as seguintes: + Adicionar a qualidade/atributo a uma mensagem de correio eletrónico e adicionar funcionalidade SCAP ao Thunderbird (<https://www.mozilla.org/en-US/thunderbird/>); + Adicionar funcionalidade SCAP ao nextcloud (<https://nextcloud.com/>) / owncloud (<https://owncloud.org/>); + Adicionar funcionalidade SCAP a software de workflow de documentos OpenKM (<https://www.openkm.com/>) ou similar ; + Adicionar funcionalidade SCAP a browser web (Chrome, Firefox e/ou Safari); + Desenvolver app mobile que coloque selo SCAP em documento; + Desenvolver comando linha (cli) de funcionalidades SCAP.

Note-se que os promotores do projeto estão abertos à adição das funcionalidades do SCAP a outras ferramentas sugeridas pelos alunos, assim como cada um dos projetos sugeridos poderá dar origem a um projeto LEI.

Para o desenvolvimento de qualquer um destes projetos, os alunos terão acesso ao ambiente de testes/qualidade do SCAP, sendo que após finalização dos testes com sucesso ficará habilitada a ser utilizada em produção.

43 SCDM (Assinatura Chave Móvel Digital) - Integração

Supervisor: jba@di.uminho.pt e jose.miranda@devisefutures.com

Contexto: Portugal, por intermédio da AMA (Agência para a Modernização Administrativa), tem desenvolvido um conjunto de projetos inovadores na área da desmaterialização de documentos e desburocratização de serviços, que se

encontram na vanguarda do que é feito a nível europeu e mundial. Um desses projetos, o SCMD, está credenciado de acordo com o regulamento UE 910/2014 (regulamento eIDAS) e permite efetuar: + Assinatura eletrónica qualificada remota, permitindo a qualquer cidadão efetuar a assinatura eletrónica de dados com uma chave privada que se encontra arquivada remotamente e, através de um dispositivo qualificado de assinatura remoto;

Objectivo: Com estes projetos pretende-se utilizar as APIs disponíveis para comunicar com o SCMD, para desenvolver add-ins/módulos/plug-ins para vários sistemas, permitindo que as funcionalidades do SCMD sejam mais facilmente acedidas nas ferramentas utilizadas pelos cidadãos no dia a dia.

Projetos: As ferramentas às quais se pretende adicionar as funcionalidades do SCMD, através de add-ins/módulos/plug-ins, são as seguintes: + Adicionar funcionalidade SCMD ao Thunderbird (<https://www.mozilla.org/en-US/thunderbird/>), tendo por objetivo assinar o correio eletrónico com SCMD; + Adicionar funcionalidade(s) SCMD ao nextcloud (<https://nextcloud.com/>) / owncloud (<https://owncloud.org/>); + Adicionar funcionalidade(s) SCMD a software de workflow de documentos OpenKM (<https://www.openkm.com/>) ou similar ; + Adicionar funcionalidade(s) SCMD a browser web (Chrome, Firefox e/ou Safari); + Desenvolver comando linha (cli) de funcionalidades SCMD.

Note-se que os promotores do projeto estão abertos à adição das funcionalidades do SCMD a outras ferramentas sugeridas pelos alunos, assim como cada um dos projetos sugeridos poderá dar origem a um projeto LEI.

Para o desenvolvimento de qualquer um destes projetos, os alunos terão acesso ao ambiente de testes/qualidade do SCMD, sendo que após finalização dos testes com sucesso ficará habilitada a ser utilizada em produção.

44 SCMD (Assinatura Chave Móvel Digital) - apps para utilizador final

Supervisor: jba@di.uminho.pt e jose.miranda@devisefutures.com

Contexto: Portugal, por intermédio da AMA (Agência para a Modernização Administrativa), tem desenvolvido um conjunto de projetos inovadores na área da desmaterialização de documentos e desburocratização de serviços, que se encontram na vanguarda do que é feito a nível europeu e mundial. Um desses projetos, o SCMD, está credenciado de acordo com o regulamento UE 910/2014 (regulamento eIDAS) e permite: + Assinatura eletrónica qualificada remota, permitindo a qualquer cidadão efetuar a assinatura eletrónica de dados com uma chave privada que se encontra arquivada remotamente e, através de um dispositivo qualificado de assinatura remoto.

Objectivo: Com estes projetos pretende-se utilizar as APIs disponíveis para comunicar com o SCMD, para desenvolver add-ins/módulos/plug-ins para vários sistemas, permitindo que as funcionalidades do SCMD sejam mais facilmente acedidas nas ferramentas utilizadas pelos cidadãos no dia a dia.

Projetos: Estes projetos têm como ponto de partida um protótipo de assinatura SCMD através de app mobile desenvolvida pela Devise Futures, em tecnologia react native.

Pretende-se (cada um dos pontos dará origem a um projeto autónomo): +

Finalizar o protótipo em tecnologia react native (ou passá-lo para outra tecnologia), transformando-o num produto final a ser disponibilizado em plataforma android; + Transformar o protótipo num produto final, utilizando tecnologia Xamarin, a ser disponibilizado em plataforma android e iOS neste caso, seria reutilizada toda a parte de assinatura SCMD, mas teria que ser refeita a parte de GUI ; + Transformar o protótipo num produto final, utilizando tecnologia Java (ou outra a determinar), a ser disponibilizado em plataforma macOS, Windows e Linux neste caso, seria reutilizada componente de assinatura SCMD, mas teria que ser refeita a parte de GUI .

Para o desenvolvimento de qualquer um destes projetos, os alunos terão acesso ao ambiente de testes/qualidade SCMD, sendo que após finalização dos testes com sucesso ficará habilitada a ser utilizada em produção.

45 Gestão de Transferências de Grandes Volumes de Tráfego em Redes WAN

Supervisor: Pedro Sousa (pns@di.uminho.pt)

Enquadramento

As Wide Area Networks (WANs) são redes que se estendem por áreas geograficamente alargadas podendo ser geridas por determinadas entidades que fornecerem serviços de conectividade Internet (e.g. ISPs). Hoje em dia proliferam sistemas tais como soluções de Cloud computing, Data-centers distribuídos em diferentes localizações da rede, sistemas de Content Distribution Networks (CDNs), entre muitos outros exemplos. Neste sentido, a necessidade constante de interligação entre os vários componentes que integram estes sistemas, origina que muitas vezes as redes ISP/WAN são atravessadas por altos volumes de tráfegos durante períodos de tempo consideráveis, tendo por vezes este tráfego também diferentes requisitos e deadlines. Poderão existir situações em que tais transferências de altos volumes de tráfego poderão causar situações de congestão na rede do ISP.

Neste contexto, pretende-se propor e desenvolver neste projeto um protótipo demonstrativo de um sistema de gestão de transferências de grandes volumes de tráfego em redes WAN. O sistema a desenvolver terá em conta os pedidos efetuados pelos clientes da rede em questão (ISPs, WAN, etc.) e, mediante as características dos pedidos efetuados (e.g. taxas de transferência pretendidas, intervalos de tempo requeridos, nível de prioridade, entre outros a definir) estabelecer critérios que organizem e despoletem essas mesmas transferências de forma otimizada de forma a não degradar a qualidade de serviço da rede. Para tal deverá ser levado em linha de conta as diferentes configurações da rede tais como, topologia, níveis de utilização dos links, configuração de routing usada e paths correspondentes entre os diversos nós da rede.

O sistema a desenvolver poderá ter como base infraestruturas configuradas com protocolos de encaminhamento IP tradicionais (e.g. OSPF ou outros), que poderão coexistir com a possível utilização de abordagens de redes Overlay que possibilitem optar por estratégias de encaminhamento alternativas, ou desviar tráfego de determinadas zonas da rede. O sistema a desenvolver poderá ser coordenado por uma entidade central, que estabelece regras para encaminhamento

de determinados fluxos específicos de tráfego recorrendo ao encaminhamento tradicional do nível da rede ou, quando necessário, recorrendo alternativamente ao nível do encaminhamento da rede Overlay. As decisões a adoptar pela entidade coordenadora poderão estar também sujeitas a determinadas políticas de engenharia de tráfego impostas pelo ISP (proteção ou priorização de determinados links/routers para as transferências em causa, limites máximos de tráfego a circular na rede, etc.).

Objectivos

Para atingir os objectivos do projeto as seguintes tarefas são definidas: i) investigação inicial nos temas relacionados com o sistema a desenvolver; ii) especificação do sistema e mecanismos a desenvolver; iii) implementação de um protótipo do sistema e iv) testes ilustrativos da operação da solução.

A linguagem de programação para implementar o protótipo é livre. Para testar o protótipo desenvolvido, deve-se no entanto assegurar que o sistema/aplicações desenvolvidas possam ser executadas e testadas num emulador de rede (sugere-se o uso do emulador Core <http://marco.uminho.pt/disciplinas/CC-MIEI/xubuncore.html>).

46 Captura e Análise de I/O de Aplicações

Supervisor: João Paulo - jtpaulo@di.uminho.pt e Tânia Esteves - tania.c.araujo@inesctec.pt

Este projeto tem como objetivo estudar e otimizar o sistema de tracing Re-Animator (<https://github.com/SNIA/reanimator-lttnng>) para capturar pedidos de rede e disco de aplicações. Estas operações serão guardadas utilizando a ferramenta Elasticsearch de forma a permitir uma análise mais detalhada. Tal análise é fundamental para perceber potenciais problemas de desempenho e funcionamento destas aplicações.

47 Análise de imagens de endoscopia - Active Learning na deteção da mucosa em VCE

Supervisor: pja@di.uminho.pt

Atualmente, a endoscopia por vídeo cápsula é um método padrão para deteção de doenças gastro-intestinais, sendo-lhe reconhecido potencial para se tornar no principal método de triagem. O diagnóstico é feito por profissionais de medicina através da visualização dos vídeos produzidos pelas cápsulas, com a duração de 8 a 10 horas, requerendo a anotação de imagens relevantes que demora tipicamente entre 60 a 90 minutos. Este processo é difícil, tedioso e propenso ao erro humano, sabendo-se que a sua performance está longe de ser perfeita. O estabelecimento de um Diagnóstico Assistido por Computador para a endoscopia por vídeo cápsula tem merecido a atenção de inúmeros grupos de investigação, no sentido de melhorar a eficiência e diminuir o esforço do diagnóstico por parte dos profissionais de medicina. Uma estratégia relevante está na redução do tempo de revisão dos vídeos, por exemplo através da ocultação de imagens "não informativas" que impedem a observação das paredes do trato intestinal. Estas imagens "não informativas" representam casos em que a mucosa é ocultada por conteúdos intestinais, como resíduos de comida, fezes, bolhas de

ar ou ainda imagens que estão saturadas, turvas ou desfocadas. Outra estratégia importante na redução do tempo de revisão dos vídeos, prende-se também com a ocultação de imagens redundantes que resultam do avanço lento da cápsula pelo trato intestinal. Neste capítulo, os avanços na aplicabilidade de algoritmos de Machine Learning, (nomeadamente DeepLearning) vieram revolucionar as técnicas convencionais, apresentando resultados significativamente melhores e extraindo características de forma automática a partir dos dados. Novas estratégias usando Active Learning têm vindo a ser usadas para aumentar ainda mais os ganhos dos algoritmos. Estas estratégias permitem envolver os profissionais de medicina na anotação de imagens selecionadas pelos algoritmos, por forma a que estas permitam melhorar o desempenho do modelos treinados para identificar características nos dados.

48 Sistema de identificação eletrónica

Supervisor: jose.miranda@devisefutures.com e jba@di.uminho.pt

O Regulamento eIDAS (Regulamento (UE) n.º 910/2014 do Parlamento Europeu e do Conselho de 23 de julho de 2014 relativo à identificação eletrónica e aos serviços de confiança para as transações eletrónicas no mercado interno e que revoga a Diretiva 1999/93/CE), no seu Capítulo II (Identificação Eletrónica), define os critérios de elegibilidade para notificação dos sistemas de identificação eletrónica, assim como os níveis de garantia dos mesmos ("reduzido", "substancial" e "elevado"). Nesse âmbito, Portugal já tem dois sistemas de identificação eletrónica aprovados, pela Cooperation Network (<https://ec.europa.eu/cefdigital/wiki/display/EIDCOMMUNITY>) com nível de garantia "elevado": + Cartão de Cidadão, publicado a 28/02/2019 ([https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52019XC0228\(01\)](https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52019XC0228(01))); + Chave Móvel Digital, publicado a 08/04/2020 (https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:JOC_2020_116_R_0005).

A ENISA (<https://www.enisa.europa.eu/>) definiu um modelo de maturidade ("A Maturity Model Framework for eID Schemes") em que define 5 níveis para avaliar a maturidade do sistema de identificação eletrónica, nas componentes de "Enrolment", "Electronic identification means management and authentication" e "Providers management and organisation".

O objetivo deste projeto é analisar o modelo de maturidade da ENISA e avaliar em que nível de maturidade é que se encontra cada um dos sistemas de identificação portugueses aprovados, assim como identificar quais as lacunas existentes para poder estar de acordo com o nível máximo (nível 5), propondo ações para que tal possa ocorrer.