

Especificação da Etapa I do Trabalho Prático

***O trabalho deve ser realizado em grupos de 3 alunos. A entrega do código, arquivos de configuração e documentação de uso deve ser realizada pelo Moodle, em um único arquivo (.zip).**

Grafos de Conhecimento para Gerenciamento de Redes

O monitoramento e a análise do tráfego de rede são cruciais para garantir o QoS (Quality of Service), detectar anomalias e otimizar o desempenho da rede. No entanto, redes modernas enfrentam desafios cada vez mais severos no gerenciamento de dados de tráfego de diferentes fontes, cada uma com seu próprio formato.

Um *knowledge-based system* é um programa que utiliza uma **base de conhecimento** e um **mecanismo de inferência** para resolver problemas complexos. Esses sistemas são projetados para simular a capacidade de resolução de problemas e tomada de decisão dos seres humanos, o que os torna ferramentas valiosas em diversos domínios—incluindo no gerenciamento de redes.

Grafos de conhecimento (*knowledge graphs* em inglês) são a estrutura de dados que armazena a base de conhecimento do sistema. Ela é baseada em [tecnologias web semânticas](#), capazes de fornecer uma abordagem estruturada para integrar, correlacionar e raciocinar sobre dados heterogêneos. Já o mecanismo de inferência é a ferramenta capaz de executar uma lógica dedutiva sobre a base de conhecimento para inferir informações que não seriam acessíveis de outras formas.

O trabalho da disciplina consiste em elaborar um *knowledge-based system* capaz de atuar com operações de gerenciamento de redes. **Para a primeira etapa, será necessário fazer a tradução de uma base de dados de informações de gerenciamento (como as MIBs do SNMP ou modelos YANG usados, por exemplo, pelo NETCONF) para o formato de grafo de conhecimento.**

O Framework de Grafos de Conhecimento

As tecnologias web semânticas foram desenvolvidas e padronizadas pelo World Wide Web Consortium (W3C) e são fortemente correlacionadas com conceitos de ontologias. Alguns componentes chave são os seguintes:

- **Resource Description Framework (RDF)**: é um modelo de dados baseado em grafos. As informações são representadas como um grafo direcionado com vértices e arestas rotuladas. A representação é feita através de triplas <sujeito, predicado, objeto>. O sujeito é um nodo do grafo, o predicado é uma aresta até o objeto, e o objeto também é um nodo do grafo. Cada uma dessas partes pode ser identificada por um *Uniform Resource Identifier* (URI).

- *RDF Schema* ([RDFS](#)): basicamente uma extensão do RDF para dar melhor suporte às ontologias. Usando RDFS é possível definir classes, subclasses e restrições, o que permite uma representação de dados mais estruturada e significativa.
- *Web Ontology Language* (OWL): é uma família de linguagens de ontologia descritiva que adiciona construções ontológicas além daquelas introduzidas pelo RDFS. Ela categoriza propriedades entre objetos e dados, e também permite adicionar restrições às propriedades, possibilitando uma representação de conhecimento mais precisa.
- *SPARQL Protocol and RDF Query Language* (SPARQL): é uma linguagem para consultar e manipular dados armazenados em formato RDF. SPARQL pode unir automaticamente todos os objetos de um grafo a partir de uma consulta, permitindo uma recuperação de dados complexa e eficiente. **Fazer consultas na base de conhecimento será a segunda etapa do trabalho.**

Etapa I do Trabalho Prático

Para a primeira etapa do trabalho prático, será necessário traduzir uma base de dados de gerenciamento de redes à sua escolha para o formato de grafo de conhecimento.

Como sugestão de base de dados de gerenciamento, modelos baseados em YANG são atualmente o foco de grupos de trabalho do [IETF](#) dentro do tópico e, portanto, sua tradução para RDF é melhor documentada. Porém, também existem trabalhos acadêmicos propondo extensões do SNMP para a tradução de MIBs.

Quanto à representação usando grafos de conhecimento, **não é obrigatório usar os padrões definidos pela W3C** apresentados na seção anterior. Existem frameworks disponíveis na internet que implementam sua própria versão de modelos de conhecimento baseados em ontologias. O [TypeDB](#), por exemplo, constrói em cima de estruturas de dados RDF mas implementa sua própria linguagem de descrição de ontologias ao invés de usar OWL.

Requisitos da Etapa 1:

1. Escolher **pelo menos 2 módulos** de uma estrutura de dados de gerenciamento de redes que esteja definida em um padrão da internet (por exemplo, dois módulos YANG diferentes, ou duas MIBs do SNMP, ou duas instâncias de algum outro protocolo). A escolha tanto do protocolo quanto dos módulos que serão usados fica à critério do aluno. Os módulos podem ser fornecidos por fabricantes de equipamentos de redes, por órgãos de padronização da Internet, entre outras fontes.
2. Para cada estrutura, fazer a tradução completa das informações presentes no módulo para uma estrutura de dados na forma de grafos de conhecimento como, por exemplo, RDF, **de forma automatizada** usando ferramentas de software.

As estruturas escolhidas e os resultados obtidos deverão ser apresentados em sala de aula na forma de uma demonstração usando a ferramenta desenvolvida pelos alunos. Deverão ser entregues todos os arquivos fonte gerados pelo grupo, bem como arquivos de configuração e algum tipo de roteiro documentado para executar as ferramentas.

Recursos Adicionais

Dentro do IETF

[Draft sobre o uso de grafos de conhecimento em operações de rede](#). A seção 9.2 fala sobre a tradução de YANG para RDF.

[Draft com alguns exemplos](#) de design de ontologias para grafos de conhecimento.

[Draft \(expirado\) especificamente sobre a construção de grafos de conhecimento usando bases de dados de gerenciamento de redes em YANG](#), com exemplos práticos de mapeamento nos apêndices. [Código no github dos autores](#) com a biblioteca de ontologias baseadas em YANG que eles começaram. [Artigo dos autores](#) sobre o mesmo assunto do draft.

Ferramentas

[RDFLib](#), um pacote em Python pra trabalhar com RDF.

[Github](#) do TypeDB. [Post no Medium](#) explicando a diferença entre o TypeDB e o OWL. [Post no blog do TypeDB](#) explicando o que é um grafo do conhecimento (com exemplos gráficos e didáticos).

Artigos

[Artigo de 2015](#) armazenando informações do Netflow em uma base de dados ontológica.

[Artigo de 2014](#) propondo uma extensão para o SNMP para dar suporte a ontologias.