Universidade
do Minho

Agentes e Sistemas Multi-Agente

Sistema de Deteção de Intrusões

- David Teixeira PG55929
- Eduardo Cunha PG55939
- Jorge Rodrigues PG55966
- Tiago Rodrigues PG56013

Maio, 2025

Introdução

Os ataques informáticos têm-se tornado cada vez mais frequentes e sofisticados, representando uma ameaça crescente para organizações de todas as dimensões. Os Sistemas de Deteção de Intrusões (IDS) tradicionais enfrentam limitações significativas, nomeadamente em termos de escalabilidade e custos operacionais.

Como resposta a este desafio, propomos um IDS distribuído, baseado em agentes inteligentes, capaz de detectar e responder de forma eficiente a intrusões, mesmo em ambientes complexos e em constante evolução.

Definição de Domínio

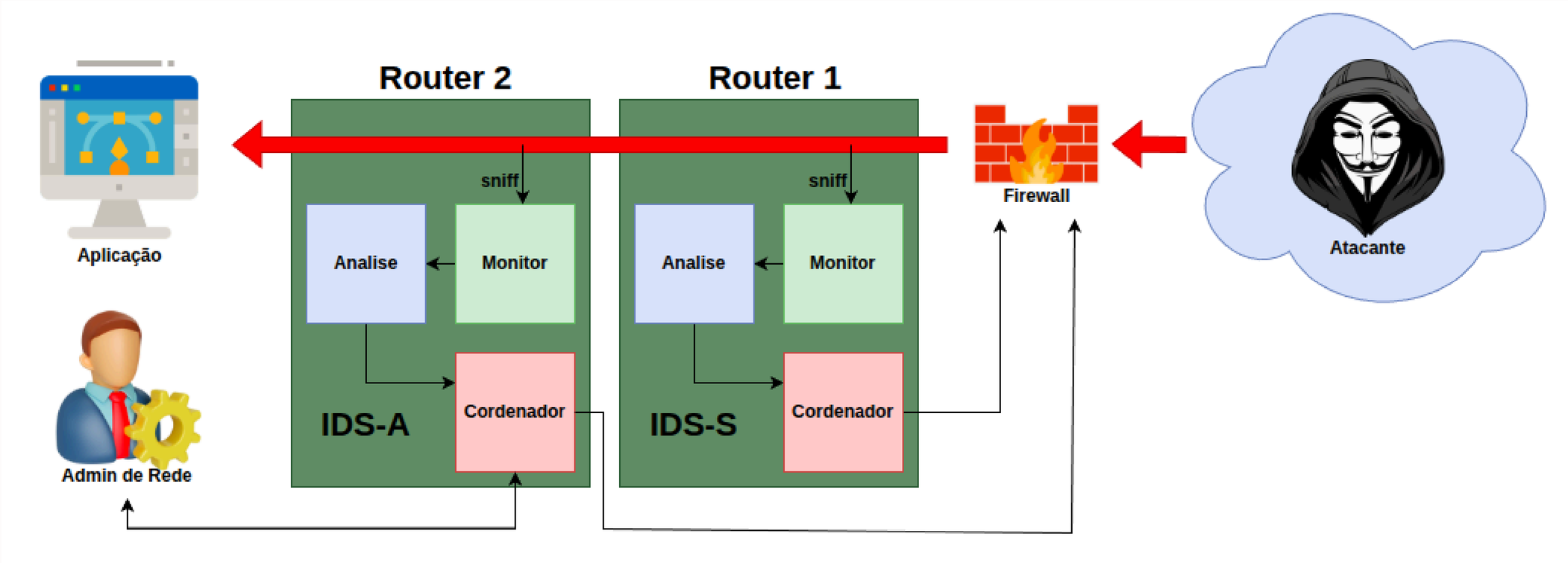
Um IDS (Sistema de Deteção de Intrusões) monitoriza o tráfego de rede com o intuito de identificar comportamentos suspeitos que possam indicar uma intrusão.

Existem dois métodos principais de deteção: por assinaturas, que identificam ataques previamente conhecidos, e por anomalias, que detetam comportamentos fora do padrão, permitindo identificar ataques desconhecidos.

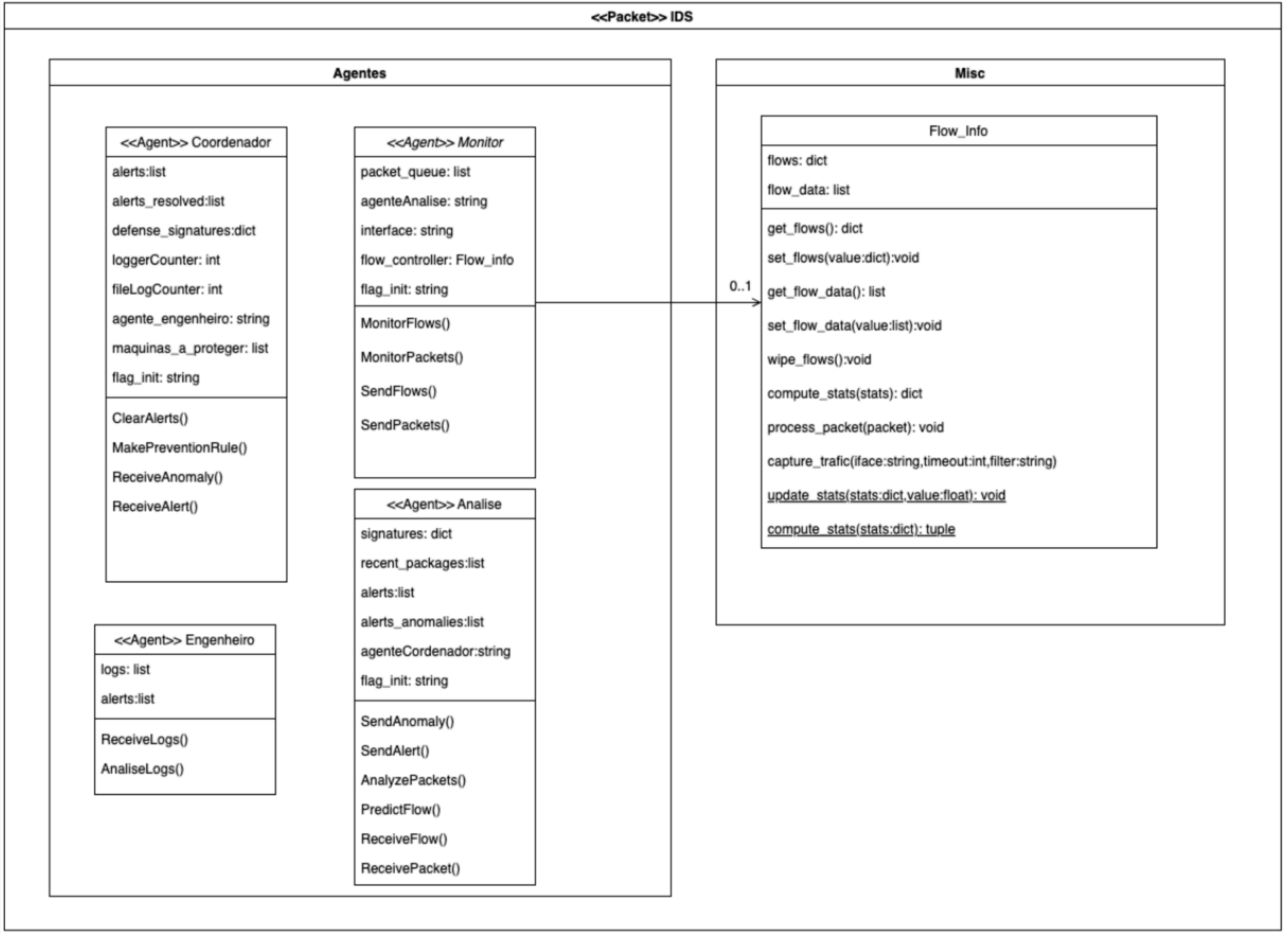
O objetivo desta proposta é desenvolver um sistema modular, facilmente configurável e com capacidade de prevenção ativa contra ameaças.

Arquitetura do Sistema

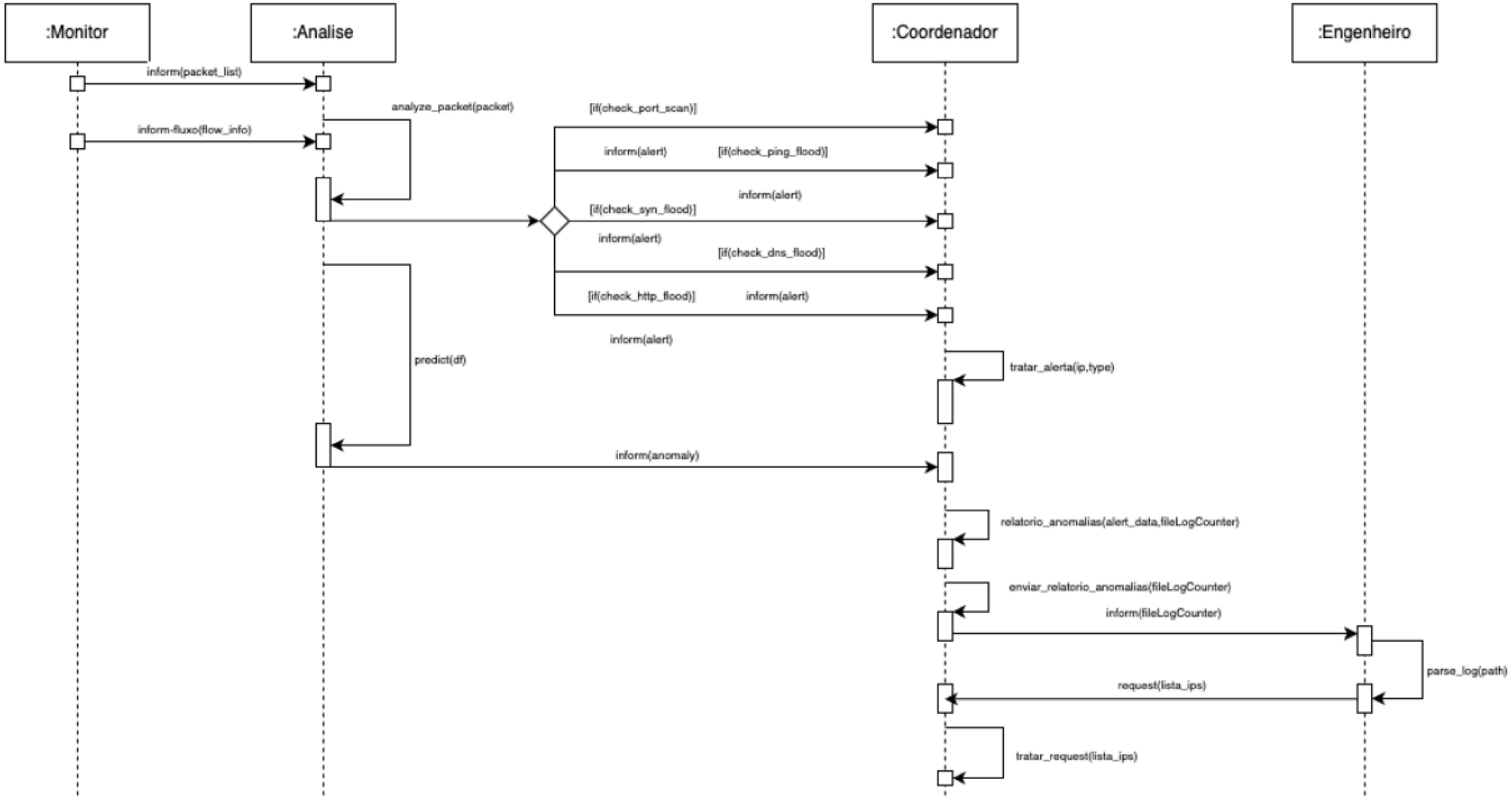
O sistema proposto é distribuído e organiza-se em duas camadas complementares de deteção: assinaturas e anomalias.
A camada de assinaturas é responsável pela deteção rápida e pelo bloqueio automático de ataques previamente conhecidos.
Já a camada de anomalias utiliza técnicas de machine learning para deteção adaptativa de comportamentos suspeitos, permitindo identificar ameaças novas ou desconhecidas.



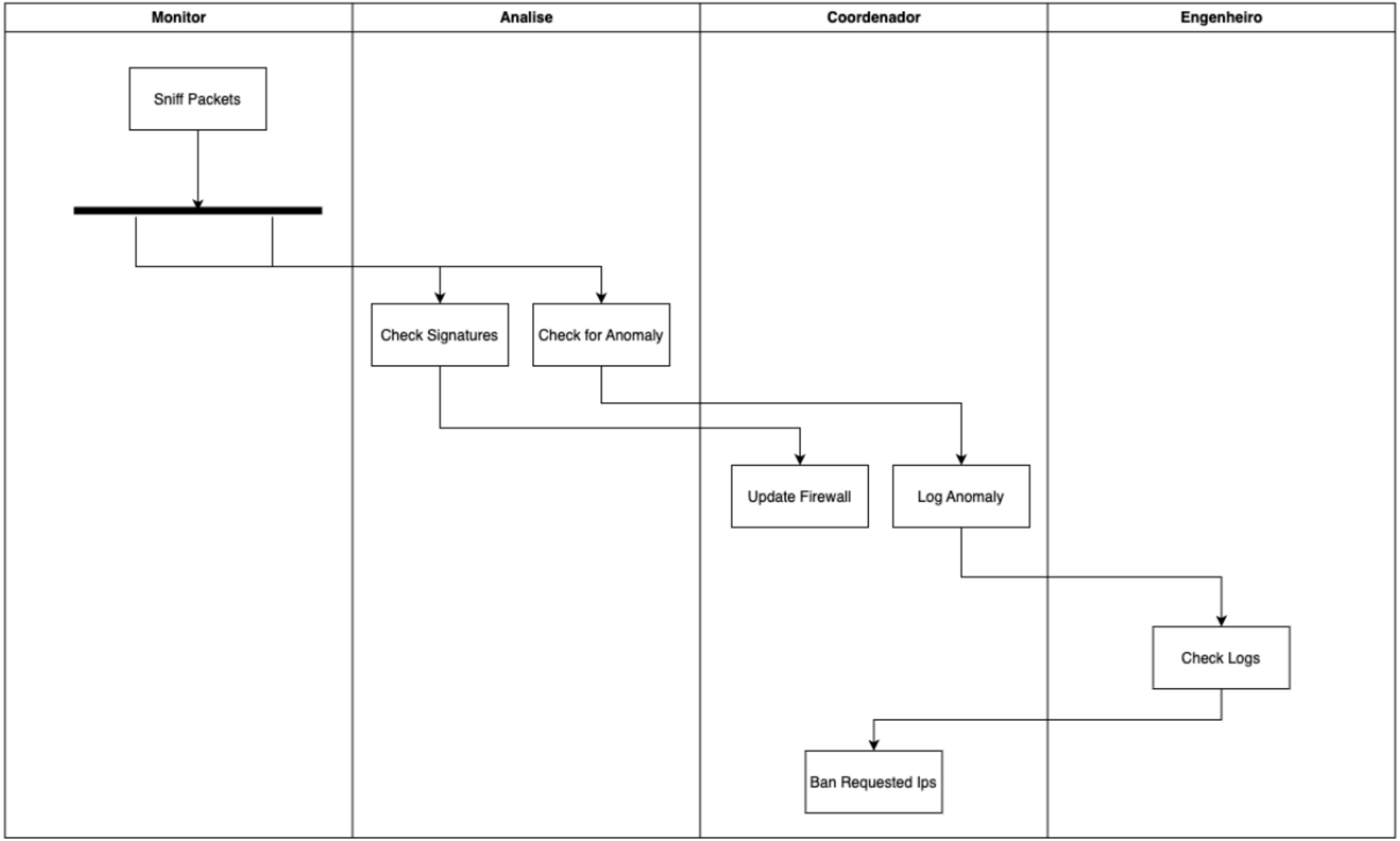
Agentes do Sistema



Funcionamento dos Agentes

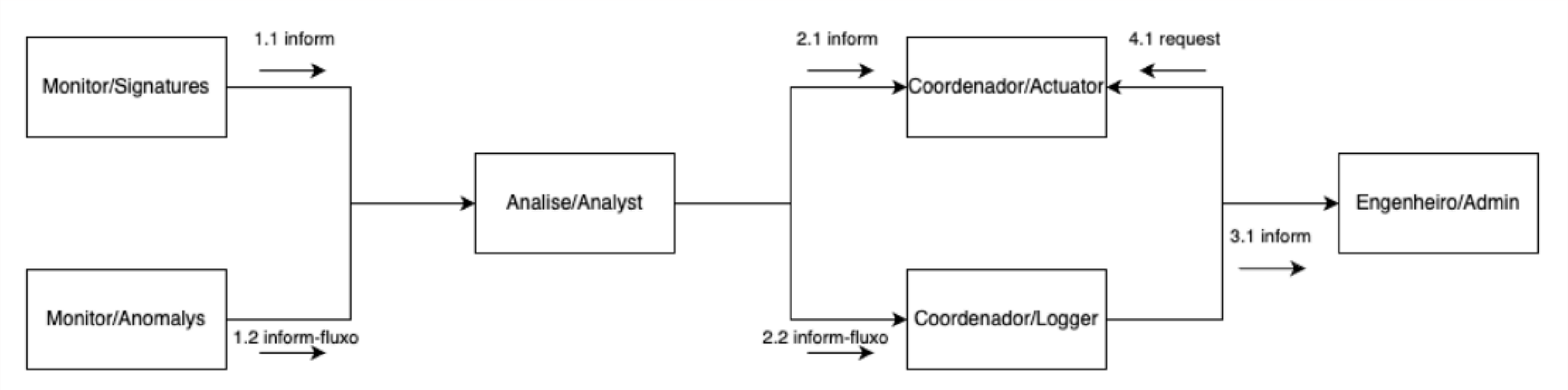


Funcionamento dos Agentes



Comunicação entre Agentes

Os agentes comunicam de forma cooperativa e hierárquica, assegurando uma coordenação eficiente das ações do sistema.
A troca de informações é feita através de mensagens dos tipos “inform” e “request”, de acordo com o padrão FIPA ACL (Agent Communication Language), promovendo interoperabilidade e clareza na comunicação.
Contudo, para aumentar a flexibilidade na utilização do IDS, nem todas as regras da FIPA ACL foram seguidas de forma estrita. Foi introduzido um novo tipo de performativa, “inform-fluxo”, adaptado às necessidades específicas do sistema.



Inteligência dos Agentes

A inteligencia dos agentes de deteção de intrusões assenta numa abordagem hibrida que combina técnicas determinísticas baseadas em regras e machine learning:

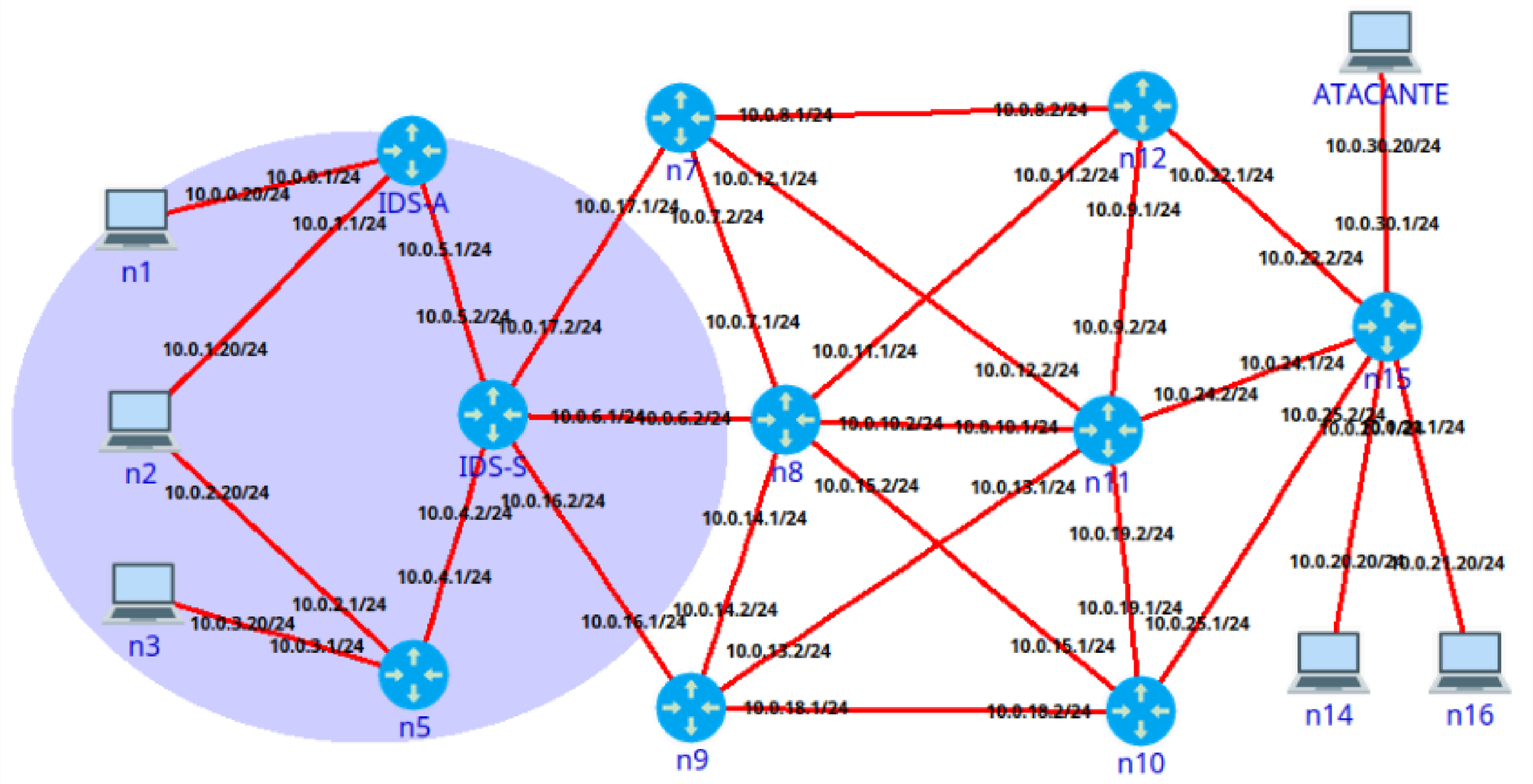
- **Assinaturas:** baseia-se em regras pré-definidas para identificar ataques conhecidos, como port scan, ping flood, entre outros.
- **Anomalias:** recorre a modelos de machine learning para detetar padrões de comportamento inesperados.

A combinação destes dois métodos aumenta a robustez e a eficácia do sistema, permitindo uma resposta mais completa face a diferentes tipos de ameaças.

```
ATTACK_SIGNATURES = {
    "port_scan": {
        "conditions": {
            "time_window": 5,  # segundos
            "min_attempts": 3
        }
    }
}

DEFENSE_SIGNATURES = {
    "port_scan": {
        "action": "block_ip",
        "description": "Bloqueia IPs que dão scan em multiplas portas num curto
intervalo de tempo.",
        "command": lambda ip: f"sudo iptables -A FORWARD -s {ip} -j DROP\n"
    }
}
```

Topologia da Rede



Demonstração

Proteção de Dados

- Análise apenas dos cabeçalhos dos pacotes, sem inspeção do conteúdo das comunicações.
- As informações analisadas são métricas agregadas e estatísticas, sem qualquer identificação de utilizadores.
- Endereços IP foram excluídos dos dados de treino dos modelos de deteção de anomalias, prevenindo critérios discriminatórios.
- Nenhum pacote é armazenado: todos são processados em tempo real e descartados após análise, utilizando o sistema de garbage collection do Python.
- Esta abordagem assegura a eliminação de riscos de retenção de dados e facilita o cumprimento do RGPD.

Trabalhos Futuros

- Expansão do sistema de assinaturas e melhoria dos modelos de machine learning, para abranger uma gama mais ampla de ameaças.
- Aumentar a escalabilidade e modularidade da arquitetura, facilitando a adaptação a diferentes ambientes de rede.
- Realizar testes mais robustos em ambientes reais ou emuladores avançados, para validar o desempenho em cenários próximos do mundo real e distribuir a hospedagem dos agentes.
- Integrar inteligência artificial agentic no Agente Engenheiro, com o objetivo de simular comportamentos e decisões semelhantes aos de um analista humano.

Conclusão

O sistema de deteção de intrusões multi-agente desenvolvido demonstra que é possível construir uma solução distribuída, cooperativa e eficaz, combinando técnicas de deteção por assinaturas e por anomalias. A arquitetura modular, baseada em agentes especializados para monitorização, análise, coordenação e engenharia, permitiu alcançar uma maior escalabilidade, flexibilidade e robustez, reduzindo a incidência de falsos positivos e assegurando a proteção dos dados dos utilizadores, em conformidade com o RGPD. Os resultados obtidos comprovam a capacidade do sistema em identificar e mitigar ataques comuns, como port scan e SYN flood, bem como a sua adaptabilidade para detetar ameaças desconhecidas através de modelos de machine learning. O sistema encontra-se preparado para evoluir, com potencial para aumentar o conjunto de assinaturas, refinar os modelos de deteção e adotar arquiteturas multinível, promovendo maior especialização e distribuição da carga computacional.

Universidade
do Minho

Agentes e Sistemas Multi-Agente

Sistema de Deteção de Intrusões

- David Teixeira PG55929
- Eduardo Cunha PG55939
- Jorge Rodrigues PG55966
- Tiago Rodrigues PG56013

Maio, 2025