# IBEMCS: IDS Baseado em Eventos Multi-Contexto para SCADA

**Anderson Mussel D'Aquino[13], Luiz Fernando Rust da Costa Carmo[12], Luci Pirmez[1], Claudio Miceli[1]**

[1]Programa de Pós-Graduação em Informática - Universidade Federal do Rio de Janeiro – Cidade Universitária – 21.941-901 – Rio de Janeiro – RJ – Brasil

[2]Instituto Nacional de Metrologia, Normalização e Qualidade Industrial – 25.250-020 – Duque de Caxias – RJ – Brasil

[3]Petrobras Transportes S.A. – Transpetro – 20.091-060 – Rio de Janeiro – RJ – Brasil

anderson.aquino@petrobras.com.br,
{rust,luci,claudiofarias}@nce.ufrj.br

***Abstract.*** *Nowadays security mechanisms used for intrusion detection in industrial automation environment use either ICT (Information and Communication Technologies) or OT (Operation Technologies) data. This work proposes an IDS that integrates ICT and OT in order to identify the ICT and OT chain of events that leaded to industrial process flaw. Measurements show that runtime and hit are linearly proportional to the data volume processed allowing a prior planning of this framework to the work environment.*

***Resumo.*** *Atualmente os mecanismos de segurança utilizados para detecção de intrusão em ambientes de automação industrial utilizam dados exclusivos TIC (Tecnologia da Informação e Comunicação) ou TO (Tecnologia da Operação). Este trabalho propõe um IDS (Intrusion Detection System) que integra informações TIC e TO para identificação das cadeias de eventos TIC e TO que ocasionaram a falha no processo industrial. Medições demonstram que o tempo de resposta e a taxa de acerto são linearmente proporcionais ao volume de dados processados permitindo um planejamento prévio da arquitetura ao ambiente de trabalho.*

## 1. Introdução

Sistemas SCADA (*Supervisory Control and Data Acquisition*) são utilizados para operar processos industriais automatizados e exercem um papel fundamental no controle das infraestruturas críticas, tais como: geração e transmissão de energia, saneamento, usinas nucleares, sistemas bélicos, dentre outros. Uma falha em tais estruturas pode causar desde graves impactos econômicos até sério comprometimento da soberania nacional. Devido à esta importância, muita pesquisa tem sido feita para ampliação dos mecanismos de segurança aplicados a sistemas SCADA. Esta pesquisa está concentrada na construção de IDS (*Intrusion Dectection System*) customizados para ambientes SCADA. As propostas destes IDS tendem a criar uma segregação nos dados processados e trabalham ou com dados puramente de automação ou com dados TIC, sem levar em consideração a relação existente entre o funcionamento do processo industrial controlado pelo sistema SCADA e a infraestrutura TIC.

 

A operação dos sistemas SCADA é feita por operadores, que são técnicos de automação com conhecimento sobre o processo industrial e treinados para utilizar as funções do sistema SCADA para garantir a execução do processo industrial, identificar anomalias e atuar de forma a levar o sistema SCADA à condição normal de operação. Ao conjunto formado pelo sistema SCADA, operadores, CLP (Controlador Lógico Programável) e instrumentação, chamamos de TO (Tecnologia da Operação).

A utilização de elementos da infraestrutura TIC é feita em todos os níveis da arquitetura SCADA. No centro de controle as funções do sistema SCADA são hospedadas em computadores e sistemas operacionais comerciais. Nos sítios remotos, os CLPs possuem interface Ethernet. A comunicação entre o centro de controle e os sites remotos é feita através de tecnologia IP (*Internet Protocol*).

Toda esta infraestrutura TIC é operada através de um NOC (*Network Operation Center*) por técnicos que não possuem conhecimento sobre a lógica do processo automatizado e utilizam ferramentas TIC que não interpretam os dados e eventos do processo automatizado. Assim como a operação da infraestrutura TIC é segregada da operação TO, a pesquisa científica para mecanismos de segurança de sistemas SCADA também adota este mesmo critério de segregação e como consequência os trabalhos propostos focam na proteção exclusiva TIC ou TO.

Esta segregação de informações entre os contextos TIC e TO gera uma lacuna por não permitir a correlação de anomalias do processo automatizado com anomalias da infraestrutura TIC. Desta forma um atacante pode conseguir infiltrar-se na infraestrutura TIC explorando uma falha do mecanismo de segurança TIC, causar um desvio de funcionalidade no processo automatizado burlando o mecanismo de segurança TO e perpetrar um ataque uma vez que não há troca de informações entre as equipes e ferramentas TIC e TO que possa identificar o ataque. Este tipo de ataque é praticado por atacantes com interesse em controlar o processo industrial. A infraestrutura TIC não é o alvo do ataque e sim o vetor utilizado para atingir o processo industrial, estuda-lo mapear suas falhas e planejar a melhor maneira de impedir o processo industrial de atingir seus objetivos durante o maior período de tempo sem ser percebido. Exemplo clássico de ataque deste tipo é o Stuxnet, que foi largamente estudado, como pode ser visto em [Falliere et al. 2011].

Este trabalho propõe um IDS para proteção do processo industrial controlado por um sistema SCADA, denominado IBEMCS (IDS Baseado em Eventos Multi-Contexto para SCADA), que estende as propostas atuais, permitindo a integração entre IDS que processam informações de contexto TIC e IDS que processam informações de contexto TO como forma ampliar o espectro de detecção destes IDS isolados através da geração de alarmes que identifiquem a anomalia TO e os eventos TIC responsáveis por tal anomalia. Este alarme, que une informações TO e TIC, permite uma análise forense mais ágil por identificar o momento e a forma como a infraestrutura TIC causou a anomalia TO. Análise de dados simulados demonstram que tanto o tempo de reposta do IDS proposto quanto a taxa de acerto são linearmente proporcionais ao volume de dados processados, permitindo um planejamento de capacidade do IDS para o ambiente em que será utilizado.

Este trabalho está organizado como segue. Na seção 2, são apresentados os trabalhos recentes relacionados a IDS para sistema SCADA. Na seção 3 é apresentada a

 

proposta detalhada. Na seção 4 é apresentado um estudo de caso. Na seção 5 são apresentados os experimentos e simulações realizados e por fim, na seção 6, a conclusão.

## 2. Trabalhos Relacionados

A grande maioria das propostas atuais para construção de IDS para sistemas SCADA utilizam dados de contextos mutuamente exclusivos TIC ou TO.

Dentre os trabalhos que utilizam dados exclusivos TIC, podemos citar [Sayegh et al. 2014] que utilizam um método estatístico para determinar o padrão de comunicação baseado nos metadados contidos nos níveis 2 e 3 de rede, [Maglaras et al. 2014] que utilizam OCSVM para traçar o perfil de comunicação dos protocolos de rede e [Ponomarev et al. 2016] que fazem um estudo de 10 classificadores estatísticos diferentes para mapeamento do padrão de trafego entre o sistema SCADA e CLP. Todos os trabalhos mencionados propuseram seus IDS através de máquina de aprendizado para identificação do padrão de comunicação de rede do sistema SCADA e geração de alarmes caso a comunicação desvie do padrão esperado.

Ainda utilizando informações exclusivas TIC podemos citar os trabalhos que utilizam colaboração de agentes distribuídos como forma de diversificação e intensificação dos mecanismos de detecção de anomalia. [Shosha et al. 2011] propõem uma arquitetura de agentes distribuída para coleta de eventos. Cada perímetro tem um concentrador de eventos local e existe um concentrador de eventos global para coordenação de todos os agentes locais. São designados agentes específicos para cada protocolo de aplicação utilizado na arquitetura SCADA. [Schuster et al. 2012] também utilizam o conceito de colaboração de agentes distribuídos, no entanto utilizam uma arquitetura descentralizada e sem agentes especializados, que gera maior trafego e causa uma maior dependência da rede. [Cruz et al. 2014] ultrapassam a fronteira dos eventos de rede providos pelos NIDS e integram os eventos de host providos pelo HIDS permitindo análise de eventos do sistema operacional, aplicações, antivírus e etc expandindo o campo de visão do IDS e permitindo uma análise de anomalia mais profunda.

Todos os trabalhos citados focam na proteção da infraestrutura TIC, diferentemente do IBEMCS que foca na proteção do processo industrial e monitora a infraestrutura TIC para impedir que um ataque consiga controlar o processo industrial.

Já no contexto TO, os trabalhos propostos utilizam mecanismos de aprendizagem para traçar o perfil de utilização do sistema SCADA através de eventos TO e identificar anomalias para geração de alarmes TO.

[Goldenberg et al. 2013] propõem uma máquina de aprendizado para mapeamento dos estados de um sistema SCADA através de um autômato finito. Esta máquina de aprendizado é utilizada para detecção de ataques através de anomalias. [Fovino et al. 2012] propõem um mecanismo para detecção de anomalias de processo através da distância Manhattan entre o estado do sistema SCADA e seus estados críticos pré mapeados. A distância entre estados é definida pela quantidade de variáveis que diferem entre os estados. [Almalawi et al. 2016] propõem uma máquina de aprendizado baseado no algoritmo de regressão *k-nearest neighbourhood* para mapeamento dos

 

estados normais do sistema SCADA e detecção de anomalia para geração de alarmes. As anomalias são estados que não se enquadram no padrão estabelecido na fase de aprendizado do algoritmo. [Carcano et al. 2011] propõem uma linguagem para descrever os estados do sistema SCADA e a partir desta formalização, identificar os estados críticos e monitorar a evolução dos estados do sistema SCADA para geração de alarmes quando tal evolução direcionar para um estado crítico.

Além de nenhum dos trabalhos citados focados em eventos TO definir um formato padrão que descreva um alarme de forma a permitir integração com sistemas externos, nenhum deles correlaciona a falha TO com a causa proveniente da infraestrutura TIC dificultando ou impossibilitando a adoção de medidas para prevenir e impedir o problema.

## 3. Proposta

A proposta deste trabalho é a criação de um IDS capaz de garantir a segurança da execução de um processo industrial contra ataques que explorem vulnerabilidades na infraestrutura TIC capazes controlar o processo industrial de forma adversa do planejado. Este IDS utiliza o mecanismo de assinatura para a identificação do ataque através do mapeamento entre as anomalias TO e a sequência de eventos TIC que causaram tal anomalia.

Uma das vantagens da identificação de anomalias através de assinaturas é a facilidade de propagação da assinatura, depois que a anomalia tem seu comportamento estudado e mapeado, devido a integração entre bases de dados públicas de vulnerabilidade, como por exemplo (CVE), (NVD) (ambas com fomento do governo norte americano), (Quickdraw SCADA IDS), e ferramentas como Snort, que são referências de tais bases de dados de vulnerabilidade na criação e disponibilização de suas respectivas assinaturas. Outra vantagem é a alta taxa de acerto, uma vez que as assinaturas são criadas para identificar comportamentos específicos da anomalia. Mecanismos de detecção capazes identificar ataques que exploram vulnerabilidades *zero-day* sofrem de problemas com taxa de acerto além de não conseguirem compartilhar informações publicamente assim como as assinaturas.

O funcionamento do IDS proposto utiliza uma base de dados de assinaturas construída por especialistas com conhecimento da infraestrutura TIC e do processo industrial que se deseja proteger, a partir de informações coletadas sobre anomalias ocorridas na planta industrial sob controle ou de bases de dados públicas, como as bases de dados TIC citadas acima, e bases de dados TO, como RISI. Cada assinatura desta base de assinaturas define uma associação entre eventos TIC e TO provenientes de IDS TIC e TO externos. As assinaturas são do tipo *condição→alarme* onde a condição representa um conjunto de restrições, baseadas nos eventos TIC e TO, necessárias para a geração de um alarme. O IDS proposto tem duas entradas de dados: Uma para recebimento de eventos TIC e outra para eventos TO. Diferente dos eventos TIC, que contam com padrões consagrados para sua representação, não há padrão para representação dos eventos TO, desta forma é proposto um formato para representação destes eventos assim como um componente conversor que pode ser utilizado caso a entrada de dados não esteja no padrão proposto.

 

Propomos o formato do evento TO como um conjunto de pares no formato *{chave=valor}*, assim como os eventos tradicionais TIC, como pode ser visto na Figura 1 que é parte de um evento gerado pelo IDS (snort) através da aplicação da Regra snort 1. Neste formato, *chave* identifica univocamente um campo dentro de um evento e *valor* é a informação associada à chave.


```
"Evento TIC" : {
        "dport-icode" : 102,
        "source-ip" : "192.168.251.1",
        "EventoTICID" : 1201502,
        "sport-itype" : 53151,
        "event-second" : 1466369924,
        "destination-ip" : "192.168.251.2",
        "signature-revision" : 0,
        "event-microsecond" : 69146,
        "blocked" : 0
        }
```


Figura 1. Representação JSON do evento TIC detectado pela Regra snort 1

No formato do evento TO proposto, três campos são definidos como obrigatórios: *ID* que é um identificador utilizado para diferenciar os eventos TO entre si, *Mensagem* que é um campo texto com uma mensagem descritiva do evento e *Timestamp* que registra o momento da geração do evento. Além destes três campos obrigatórios, o evento TO contém outros campos, preenchidos pelo IDS TO, que transmitem informações sobre o estado do sistema SCADA relacionadas ao evento. A estrutura de um evento TO pode ser visualizada na Tabela 1.

Tabela 1. Visão de um Evento TO.

| | |
|---|---|
| ID | Tipo numérico. Chave primária. |
| Mensagem | Tipo texto. Descrição do evento. |
| Timestamp | Tipo data/hora. Registro do momento do evento. |
| Campo 1 | Valor proveniente do IDS TO. |
| Campo 2 | Valor proveniente do IDS TO. |
| ⋮ | ⋮ |
| Campo N | Valor proveniente do IDS TO. |

A arquitetura conta ainda com um sistema de decisão, que é responsável por orquestrar o funcionamento dos demais componentes para junção das informações TIC e TO e posterior geração do alarme.

## 3.1. Arquitetura lógica

A Figura 2 ilustra a arquitetura lógica do IDS, que é composta de um conjunto de componentes e uma base de dados. Os componentes são: **Conversor de Eventos TO**, **Gestor de Eventos TO**, **Gestor de Eventos TIC**, **Gestor de Decisão**, **Correlacionador** e **Gestor de Alarmes**. A base de dados é a **Base de Assinaturas**.

 

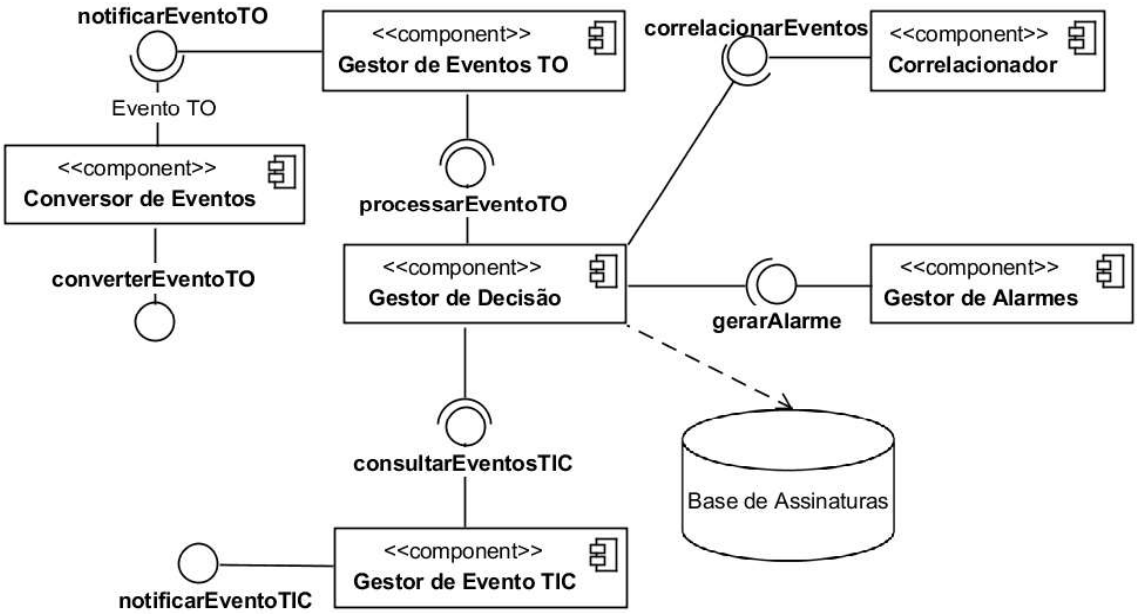


Figura 2. Arquitetura IBEMCS.

## 3.2. Base de Assinaturas

O IDS proposto utiliza o mecanismo tradicional de assinaturas onde os alarmes são gerados quando um conjunto de restrições é satisfeito. Como um exemplo de regra podemos citar a Regra snort 1 para identificação e registro de um evento de acesso de rede a um CLP na porta 102 TCP.

```
log tcp any any -> 192.168.251.2 102 ( sid:1201502; msg: "Siemens Step7 conexão de manutenção estabelecida"; )
```

Regra snort 1. Regra para geração de alarme TIC

As assinaturas podem ser entendidas como a expressão lógica *condição→alarme* armazenadas na **Base de Assinaturas** que são consultadas pelo componente **Gestor de Decisão**. Os elementos da condição necessários para geração do alarme são:

- **EventoTIC$_{ID}$**: Este campo representa o ID do evento TIC gerado pelo IDS TIC que deve ser condicionado ao Evento TO para geração do alarme.
- **EventoTO$_{ID}$**: Este campo representa o ID do evento TO gerado pelo IDS TO que deve ser condicionado ao Evento TIC para geração do alarme. Diferentemente dos Eventos TIC, não há padrão estabelecido para representação dos eventos TO, portanto, adotaremos o formato proposto neste trabalho.
- **TimeWindow**: Este campo representa uma restrição de localidade temporal entre os eventos TO e TIC mapeados na condição da assinatura. *TimeWindow* é um número do tipo real que representa a máxima diferença de tempo, em segundos, decorrido entre a geração dos eventos TO e TIC para validação de associação entre esses eventos. Este campo é fornecido pelo componente **Gestor de Decisão** como filtro para o componente **Gestor de Eventos TIC** durante a chamada do método *consultarEventosTIC*.
- **ValidFunc**: Este campo representa uma restrição de localidade espacial entre os eventos TO e TIC mapeados na condição da assinatura. *ValidFunc* é uma função

 

de validação que deve ser programada em linguagem de programação pelo especialista que desenvolveu a assinatura com a finalidade de garantir que os eventos TIC e TO estão relacionados aos mesmos elementos de rede da planta industrial. O componente **Correlacionador** avalia a função de validação com os dados dos eventos TO e TIC fornecidos pelo componente **Gestor de Decisão**.

O parâmetro *TimeWindow* influencia diretamente no poder de processamento necessário para a utilização do IDS, pois controla o volume de dados que o IDS irá processar em cada requisição de correlaciomento de eventos. O ajuste desde parâmetro depende de estudo do especialista pela implantação do IDS, que deve levar em consideração o perfil de vazão de geração de eventos da planta industrial para o correto dimensionamento dos recursos para implantação do IDS.

*ValidFunc* é responsável pela validação da correlação entre os eventos através de informações contidas nos próprios eventos. A complexidade de tais validações influencia no tempo de processamento do IDS.

O alarme da assinatura conterá as informações TIC e TO necessárias que permitam aos técnicos de segurança analisarem e tratarem o alarme. Essas informações são únicas de cada alarme e especificadas pelo especialista responsável pelo planejamento do alarme.

Uma assinatura pode ser descrita como $(EventoTIC_{ID}, EventoTO_{ID}, TimeWindow, ValidFunc) \rightarrow (CampoTIC_1, ..., CampoTIC_N, CampoTO_1, ..., CampoTO_N)$, onde $CampoTIC_i$ é um dos campos contidos no evento TIC mapeado na condição da assinatura do alarme e $CampoTO_i$ é um dos campos contidos no evento TO mapeado na condição da assinatura do alarme.

## 3.3. Conversor de Eventos TO

O funcionamento do IBEMCS depende de uma entrada de dados padronizada, tanto de eventos TIC quanto de eventos TO. Este componente é um padrão de projeto *Adapter* para uso do IDS TO que não seja capaz de gerar eventos no formato Evento TO. Este componente define as interfaces que devem ser implementadas para geração de eventos em formato evento TO. A implementação destas interfaces é especifica para cada sistema que forneça dados para este componente.

## 3.4. Gestor de Eventos TO

Este componente é responsável por gerenciar o recebimento concorrente de eventos TO e acionar o componente Gestor de Decisão para iniciar o processo de análise de dados para geração do alarme.

## 3.5. Gestor de Decisão

Este componente é responsável pelo recebimento da entrada de dados, Eventos TO, e obtenção de dados necessários para a correlação de eventos realizado pelo componente **Correlacionador**. Seu funcionamento pode ser descrito como a seguinte sequência de ações, visualizada na Figura 3:

1. Recebimento do Evento TO;

 

2. Consulta à **Base de Assinaturas** para obtenção das assinaturas relacionadas ao Evento TO recebido;
3. Consulta ao componente **Gestor de Eventos TIC** para obtenção da lista de Eventos TIC que deve ser processada pelo componente **Correlacionador**. Como filtro de pesquisa, o componente **Gestor de Eventos TIC** recebe os parâmetros *TimeWindow* e $EventoTIC_{ID}$ presentes na assinatura obtida no passo 2;
4. Acionamento do componente **Correlacionador**, informando o Evento TO obtido no passo 1, a assinatura obtida no passo 2 e a lista de Eventos TIC obtida no passo 3.

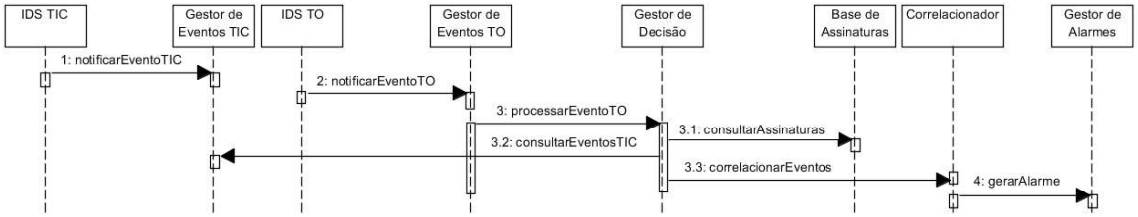


Figura 3. Diagrama de sequência do IBEMCS.

## 3.6. Gestor de Eventos TIC

Este componente é responsável por responder às solicitações do componente **Gestor de Decisão** para obtenção de Eventos TIC segundo os filtros presentes nas assinaturas da **Base de Assinaturas**. A consulta é feita através da interface *consultarEventosTIC* cujos parâmetros são *TimeWindow* e $EventoTIC_{ID}$ que indicam respectivamente a janela de tempo a ser considerada para pesquisa e o identificador do Evento TIC que de ser pesquisado.

## 3.7. Correlacionador

A função deste componente é avaliar os dados contidos nos Evento TO e Evento TIC para validar a geração do alarme. Esta avaliação é feita através da função de validação *ValidFunc* contida na assinatura do alarme. O objetivo desta função é garantir que o Evento TIC que está sendo avaliado para geração do alarme de fato afetou os componentes de automação envolvidos na geração do Evento TO que está sendo analisado. Caso este teste não seja feito poderíamos gerar um alarme a partir de eventos TO e TIC que aconteceram em momentos próximos, mas em localidades diferentes, como por exemplo o Evento TO acontecido no Brasil e o Evento TIC acontecido na Espanha. Caso a topologia de rede do sistema SCADA sendo monitorado permita que eventos ocorridos em regiões geográficas distintas tornem válida a geração de um alarme, a função de validação *ValidFunc* deve processar dados adicionais de Evento TO e/ou Evento TIC para tornar tal lógica válida.

Como exemplo, podemos definir uma função de validação espacial chamada *ValidarIP*, definida no Algoritmo 1, que realiza a validação espacial entre os eventos TIC e TO baseado nas informações de rede do dispositivo que gerou o evento.

Algoritmo 1. Algoritmo para validação da localidade espacial entre os Eventos TO e TIC

```
1.  Função ValidarIP(EventoTIC, EventoTO)
```

 

```
2.    Se EventoTIC.IP == EventoTO.CLPAddress
3.       Então
4.          retorne Verdadeiro
5.       Senão
6.          retorne Falso
```

## 3.8. Gestor de Alarmes

Este componente é responsável pela geração de um alarme como descrito na assinatura cadastrada na **Base de Assinaturas**. O componente **Correlacionador**, após executar todas as validações necessárias para geração do alarme, invoca o componente **Gestor de Alarmes** fornecendo a assinatura e os Eventos TIC e TO que foram correlacionados para geração do alarme. O componente **Gestor de Alarmes** identifica no componente alarme da assinatura, os campos dos eventos TIC e TO que devem ser preenchidos no alarme.

## 4. Estudo de caso

A **Figura 4** ilustra parte de um processo industrial controlado por um sistema SCADA e executado no centro de controle de uma Transportadora de Gás Natural. Neste exemplo estão representados o centro de controle, onde está situado o sistema SCADA, e um dos muitos sítios remotos onde ocorre parte do processo industrial. O sítio representa um trecho de um gasoduto com três válvulas controladas por um mesmo CLP. O trecho representado no exemplo não contém nenhuma junção ou derivação, portanto os valores medidos de pressão e vazão nos sensores das três válvulas devem coincidir.

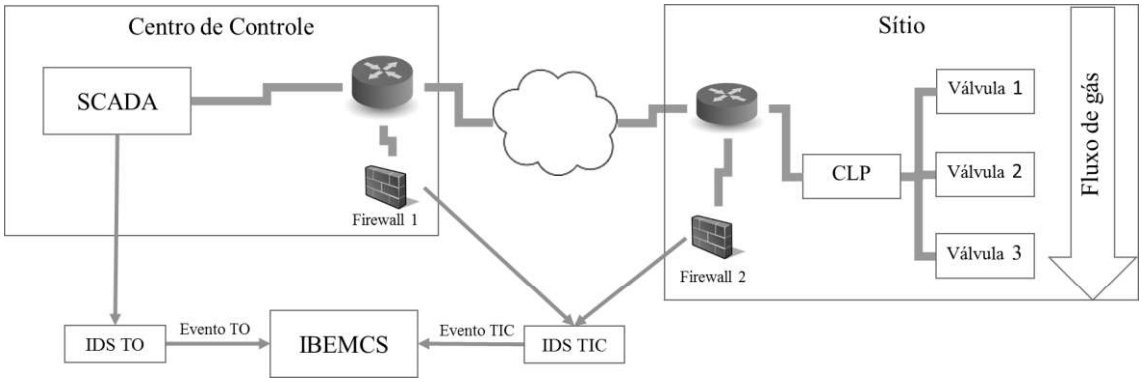


Figura 4: Estudo de caso em um processo industrial.

O acesso ao CLP é feito remotamente tanto pelo sistema SCADA para realização das tarefas de supervisão e controle executadas no centro de controle com finalidade de operação quanto por sistemas de suporte executados pela equipe de manutenção no próprio sítio. Os dados a seguir são baseados em dados de supervisão de uma Transportadora de Gás Natural.

Em um instante $t_0$ a visão do operador, no centro de controle, das variáveis pressão e vazão das três válvulas pode ser resumida pela Tabela 2.

Tabela 2. Situação normal de operação.

| | Pressão (Kgf/cm$^2$) | Vazão (m$^3$/dia) |
|---|---|---|
| Válvula 1 | 77,7 | 2.721.700 |
| Válvula 2 | 77,7 | 2.721.700 |
| Válvula 3 | 77,7 | 2.721.700 |

 

Em um instante $t_2>t_0$ há uma alteração no estado do processo automatizado e o operador a passa visualizar os valores de pressão e vazão representados na Tabela 3.

Tabela 3. Situação anormal de operação.

| | Pressão (Kgf/cm$^2$) | Vazão (m$^3$/dia) |
|---|---|---|
| Válvula 1 | 77,7 | 2.721.700 |
| Válvula 2 | 71,3 | 2.962.683 |
| Válvula 3 | 77,7 | 2.721.700 |

Um Evento TO é gerado e registrado no IBEMCS devido ao estado anormal do sistema SCADA. A configuração deste Evento TO pode ser visualizada na Tabela 4.

Tabela 4. Evento TO que representa estado SCADA anormal.

| Campo | Valor |
|---|---|
| ID | 1001 |
| Mensagem | Inconsistência de valor em linha. |
| Timestamp | 01/01/2015 17:23:31.100 |
| V1Pressao | 77,7 |
| V2Pressao | 71,3 |
| V3Pressao | 77,7 |
| V1Vazao | 2.721.700 |
| V2Vazao | 2.962.683 |
| V3Vazao | 2.721.700 |
| CLPAddress | 172.15.52.1 |

A alteração dos valores de pressão e vazão na válvula 2 é um comportamento não esperado e requer análise do operador TO. Diante de tal anomalia, o operador deve tomar uma ação para corrigir o estado do sistema SCADA, que pode ser:

- Enviar comando ao CLP para equilibrar as três válvulas;
- Acionar a equipe de manutenção do campo para inspecionar a válvula 2 e o CLP;
- Acionar a equipe de Engenharia de automação para inspecionar o sistema SCADA.

A anomalia é tratada sob o ponto de vista exclusivo de automação. As equipes envolvidas foram de automação e os dados coletados para análise da anomalia são de automação.

Em um instante $t_1<t_2$, $t_1>t_0$ acontece um evento percebido pela equipe TIC. Um acesso remoto ao CLP através da VPN de manutenção é realizado e registrado nas ferramentas de monitoração da infraestrutura TIC. Este acesso, no entanto, não chama atenção da equipe de segurança TIC pois é uma operação prevista e tem seu controle de acesso liberado.

O IDS TIC está planejado para registrar os eventos de manutenção no CLP e por isso gera um Evento TIC que é registrado no IBEMCS. A configuração deste Evento TIC pode ser visualizada na Tabela 5.

Tabela 5. Evento TIC de manutenção no CLP.

| Campo | Valor |
|---|---|
| ID | 4365 |
| Mensagem | Acesso TCP na porta de manutenção. |
| Timestamp | 01/01/2015 17:21:57.085 |
| IP | 172.15.52.1 |
| Porta | 502 |

A análise dos eventos ocorridos no processo automatizado e na infraestrutura TIC de forma isolada impossibilita a criação de nexo causal entre tais eventos e cria uma fragilidade que pode ser explorada de forma maliciosa contra o processo automatizado.

 

O IBEMCS provê os meios para preencher esta lacuna de comunicação entre os técnicos de automação e os técnicos de TIC através da geração de alarmes que contenham informações TIC e TO correlacionadas à anomalia.

Neste caso em estudo, para que a anomalia seja identificada, a Assinatura 1 deve ser cadastrada na Base de Assinaturas do IBEMCS. A condição da assinatura correlaciona o Evento TIC de ID 4365, acesso ao CLP, com o Evento TO de ID 1001, inconsistência de valores das três válvulas em sequência, em uma janela de tempo de 120s. Para validação espacial, utilizamos a função *ValidarIP* descrita no Algoritmo 1.

Assinatura 1. Assinatura para geração de alarme que correlaciona inconsistência de medição nas válvulas e acesso de manutenção no CLP.

(4365, 1001, 120, ValidarIP)→(EventoTIC.ID, EventoTIC.Mensagem, EventoTIC.IP, EventoTIC.Porta, EventoTO.ID, EventoTO.Mensagem, EventoTO. Timestamp, EventoTO.V2Pressao, EventoTO.V2Vazao)

## 5. Dados experimentais

No âmbito do presente trabalho, o IDS proposto foi implementado na linguagem Python. Para criação do ambiente de simulação foi desenvolvido um aplicativo auxiliar cuja finalidade é controlar a quantidade de eventos disponibilizados para processamento pelo IDS. Este aplicativo é multithread e a quantidade de threads é controlada em cada ensaio para controlar a vazão de dados desejada. Cada thread registra um número fixo de 4 eventos no IBEMCS, sendo 3 eventos TIC e 1 evento TO.

Os experimentos foram conduzidos com uma variação do número de threads de 1 a 400, em intervalos de 10, refletindo em uma variação na quantidade eventos simultâneos de 4 a 1600, em intervalos de 40. Cada experimento foi repetido 30 vezes, para obter um intervalo de confiança de 95%.

Os eventos TIC considerados para análise são eventos de rede gerados pelo IDS Snort, como resultado da aplicação da Regra snort 1. O evento TO selecionado foi o evento descrito na Tabela 4 do estudo de caso apresentado. A ferramenta (Elasticsearch) é utilizada para gerenciar a base de dados de eventos que foi populada com 6 milhões de registros em um período de 30 dias durante a realização dos testes. Não houve descarte de eventos.

As métricas utilizadas na avaliação são **Tempo de Resposta** e **Taxa de Acerto**. Tempo de resposta é definido como o tempo decorrido desde o recebimento do evento TO pelo IBEMCS até a geração do alarme. Taxa de Acerto é definido como a relação entre a quantidade de alarmes gerados (verdadeiros positivos) e a quantidade de anomalias processadas.

O IBEMCS, o aplicativo auxiliar gerador de eventos e a base de dados (Elasticseach) são executados em uma máquina virtual Ubuntu Linux 14.04 com dois processadores 2.30Ghz e 2.0GB RAM.

### 5.1. Tempo de Resposta

O estudo do tempo de resposta dos sistemas de segurança é fundamental para determinar a adequação de tais sistemas aos cenários onde serão utilizados. Este experimento foi

 

conduzido para estudar a evolução do tempo do IDS proposto em função do volume de dados processados.

Foram coletados dados para *TimeWindow* de 2,5s, 5,0s e 7,5s [Aramaki 2015]. O parâmetro *TimeWindow* influencia na quantidade de eventos TIC que será correlacionada ao evento TO que gerou o processo de análise.

A análise dos dados mostra que os tempos médio e máximo de resposta (Figura 5 e Figura 6, respectivamente) são lineares em função do volume de dados processados e não houve influência perceptível de *TimeWindow* nos cenários apresentados.

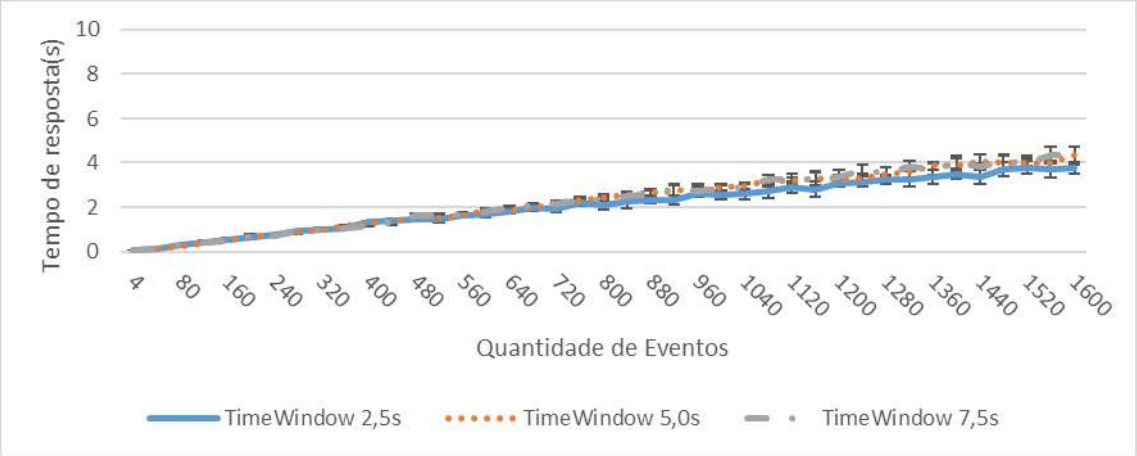


Figura 5. Tempo médio de resposta em função do número de eventos

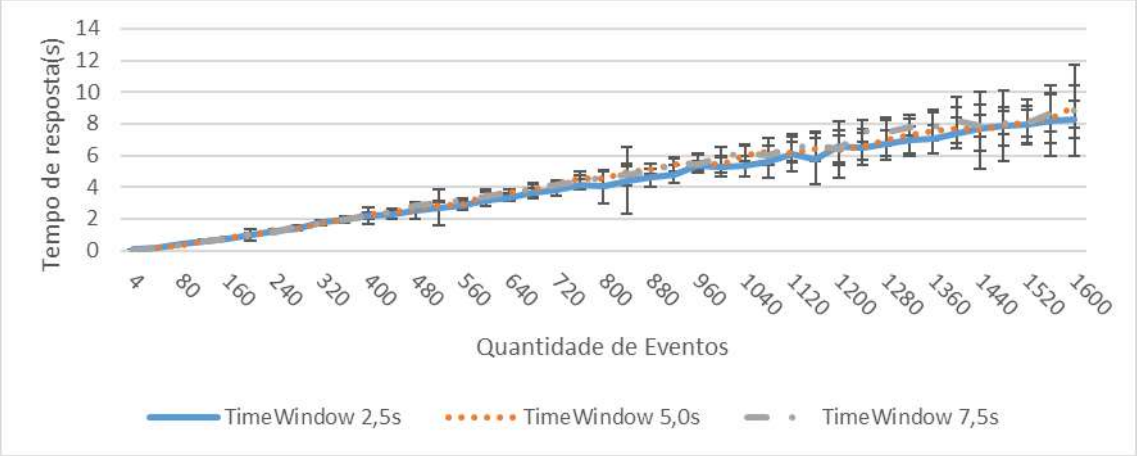


Figura 6. Tempo máximo de resposta em função do número de eventos

## 5.2. Taxa de Acerto

A Figura 7 mostra o gráfico da taxa de acerto do IBEMCS em função do número de eventos processados. Através do gráfico é possível perceber que a quantidade de eventos tem influência direta sobre a taxa de acerto. O aumento do número de eventos causa um aumento no tempo de resposta do IBEMCS. Este aumento no tempo de resposta tem impacto direto na taxa de acerto pois causa um deslocamento da janela de tempo considerada para correlacionamento de eventos fazendo com que os eventos TIC

 

percam a condição de associação aos eventos TO determinados pelo parâmetro *TimeWindow*.

O número de eventos que começa a afetar a taxa de acerto é 400, 1000 e 1400 para *TimeWindow* de 2,5s, 5,0s e 7,5s respectivamente, pois nestes limites o tempo máximo de resposta aproxima-se do valor *TimeWindow* tendo início o declínio da taxa de acerto do IBEMCS. A partir destes limites o aumento do número de eventos tem efeito inversamente proporcional sobre a taxa de acerto.

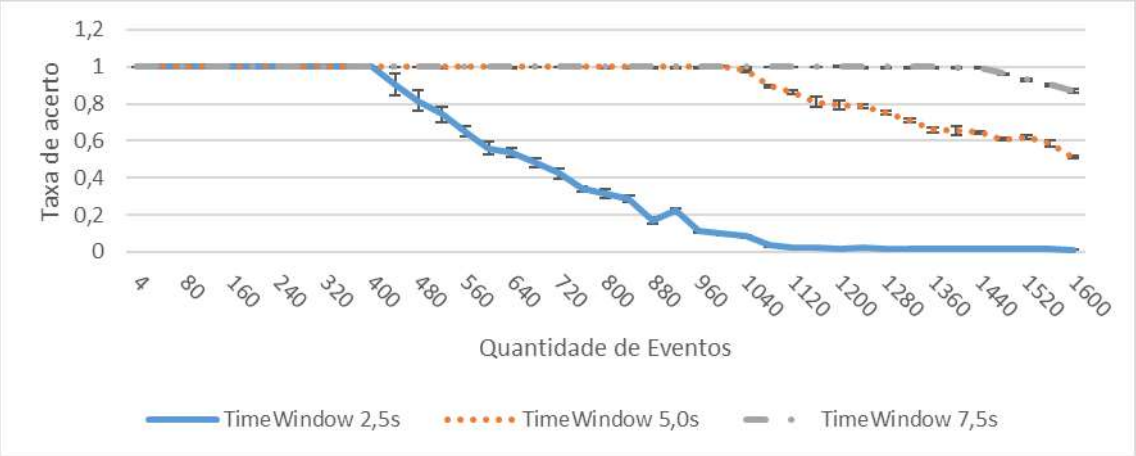


Figura 7. Taxa de acertos em função do número de eventos

## 6. Conclusão

A segurança das infraestruturas críticas depende dos mecanismos de segurança empregados para controle da lógica do processo automatizado e também da infraestrutura TIC. Atualmente existe uma lacuna nos mecanismos de segurança propostos para infraestruturas críticas. Esta lacuna é causada pela visão mutuamente exclusiva entre TIC e TO e impede correlacionar eventos TIC com TO como forma de enriquecimento de informações para aumento da acurácia dos mecanismos de detecção de intrusão. Este trabalho propõe um IDS que permite o correlacionamento de informações TIC e TO para construção de mecanismos de detecção de intrusão. O estudo do desempenho da arquitetura proposta demonstra que ela é fácil de ser ajustar a capacidade do sistema a ser monitorado pois o tempo de resposta é uma função linear do volume de informação demandada o que permite planejar também a taxa de acerto do mecanismo pois esta pode ser determinada em função do volume de informações processadas.

## 7. Referências

Almalawi, A., Fahad, A., Tari, Z., Alamri, A., Alghamdi, R., & Zomaya, A. Y. (2016). An Efficient Data-Driven Clustering Technique to Detect Attacks in SCADA Systems. IEEE Transactions on Information Forensics and Security, 11(5), 893–906. doi:10.1109/TIFS.2015.2512522

Carcano, A., Coletta, A., Guglielmi, M., Masera, M., Fovino, I. N., & Trombetta, A. (2011). A Multidimensional Critical State Analysis for Detecting Intrusions in

 

SCADA Systems. Industrial Informatics, IEEE Transactions on, 7(2), 179–186. doi:10.1109/TII.2010.2099234

Goldenberg, N., & Wool, A. (2013). Accurate modeling of Modbus/TCP for intrusion detection in SCADA systems. International Journal of Critical Infrastructure Protection, 6(2), 63–75. doi:10.1016/j.ijcip.2013.05.001

Maglaras, L. A., & Jiang, J. (2014). Intrusion detection in SCADA systems using machine learning techniques. In 2014 Science and Information Conference (pp. 626–631). doi:10.1109/SAI.2014.6918252

Nai Fovino, I., Coletta, A., Carcano, A., & Masera, M. (2012). Critical state-based filtering system for securing SCADA network protocols. IEEE Transactions on Industrial Electronics, 59(10), 3943–3950. doi:10.1109/TIE.2011.2181132

Sayegh, N., Elhajj, I. H., Kayssi, A., & Chehab, A. (2014). SCADA Intrusion Detection System based on temporal behavior of frequent patterns. MELECON 2014 - 2014 17th IEEE Mediterranean Electrotechnical Conference, (April), 432–438. doi:10.1109/MELCON.2014.6820573

Schuster, F., & Paul, A. (2012). A distributed intrusion detection system for industrial automation networks. In Proceedings of 2012 IEEE 17th International Conference on Emerging Technologies & Factory Automation (ETFA 2012) (pp. 1–4). doi:10.1109/ETFA.2012.6489703

Shosha, A. F., Gladyshev, P., Wu, S. S., & Liu, C. C. (2011). Detecting cyber intrusions in SCADA networks using multi-agent collaboration. In 2011 16th International Conference on Intelligent System Applications to Power Systems, ISAP 2011 (pp. 1–7). doi:10.1109/ISAP.2011.6082170

Tiago Cruz, Jorge Proença, Paulo Simões, Matthieu Aubigny, Moussa Ouedraogo, Antonio Graziano, L. Y. (2014). Improving Cyber Security Awareness on Industrial Control Systems: The CockpitCI Approach. 13th European Conference on Cyber Warfare and Security The University of Piraeus Greece 3-4 July 2014, (July), 326.

Snort.org. (2016). Snort. Retrieved June 24, 2016, from https://www.snort.org/

Elasticsearch. (2016). Retrieved June 24, 2016, from https://www.elastic.co/

Quickdraw SCADA IDS (2016). Retrieved September 12, 2016, http://www.digitalbond.com/tools/quickdraw/

Aramaki, T. L., Vellasco, M. M. B. R., & Barbosa, C. R. H. (2015). A neural network approach for leak detection and localization in liquid pipelines. In *8th Brazilian Congress on Metrology* (pp. 2–5). Bento Gonçalves.

NVD (2016). Retrieved September 12, 2016, https://nvd.nist.gov/

CVE (2016). Retrieved September 12, 2016, https://cve.mitre.org/

RISI (2016). Retrieved September 12, 2016, http://www.risidata.com/

Falliere, N., Murchu, L. O., & Chien, E. (2011). W32.Stuxnet Dossier. Symantec-Security Response, Version 1.(February 2011), 1–69. http://doi.org/20 September 2015

 