

ANÁLISE QUANTITATIVA DE EVENTOS CRIMINAIS UTILIZANDO ABORDAGEM SEMÂNTICA

Quantitative Analysis of Criminal Events Using Semantic Approach

DIXIM

ANÁLISE QUANTITATIVA DE EVENTOS CRIMINAIS UTILIZANDO ABORDAGEM SEMÂNTICA

Quantitative Analysis of Criminal Events Using Semantic Approach

Gustavo Marttos Cáceres Pereira¹, João Henrique Martins², Leonardo Castro Botega²

(1) Grupo de Interação Humano-Computador, Centro Universitário Eurípides de Marília, Av. Hygino Muzzi, Marília – SP

marttos@univem.edu.br

(2) Stratelli – Inteligência Estratégica, Av. Hygino Muzzi, Marília – SP

joao.henriq.75@gmail.com, botega@univem.edu.br

Resumo:

O gerenciamento de informações de riscos utilizando dados criminais apresenta desafios associados à aquisição de Consciência Situacional, tais como a dinamicidade, heterogeneidade, variedade e o grande volume de dados. Além disso, representar adequadamente as informações no domínio criminal, tais como dados sobre vítimas, criminosos, locais e a própria situação de crime, pode contribuir para processos de análise quantitativa de dados, tornando assim a tomada de decisão ser mais assertiva. Ao mesmo tempo, há ainda a necessidade do desenvolvimento de modelos semânticos que representem a realidade das situações de ambientes de riscos, o que produz ainda mais desafios para a quantificação de entidades e características relevantes. Portanto, este trabalho tem por objetivo apresentar o processo de desenvolvimento de quantificação de dados em ontologias para o domínio de gerenciamento de riscos, visando suportar a extração de dados específicos para ampliar a capacidade de obtenção de Consciência Situacional e permitir que as decisões tomadas, baseadas nestas informações, sejam mais assertivas. Para tal, será empregada a análise de tarefas dirigidas por objetivos, a análise de vocabulários e propriedades no contexto criminal a fim de restringir sua interpretação e a avaliação de pesos e valores semânticos desses vocabulários. Sendo assim, ao final, será possível mensurar as informações necessárias dentro de um sistema de gerenciamento de riscos.

Palavras-chave: Consciência Situacional; Ontologia; Análise Quantitativa; Gerenciamento de Riscos

Abstract:

The management of risk information using criminal data presents challenges associated with the acquisition of Situational Awareness, such as dynamicity, heterogeneity, variety and large data volume. In addition, adequately representing information in the criminal domain, such as data on victims, criminals, places and the actual crime situation, can contribute to quantitative data analysis processes, thus making decision making more assertive. At the same time, there is still a need for the development of semantic models that represent the reality of situations in risk environments, which produces still more challenges for the quantification of entities and relevant characteristics. Therefore, this paper aims to present the process of a development of data quantification in ontologies for the domain of risk management, aiming at support the extraction of specific data to increase the capacity to obtain Situational Awareness and allow the decisions taken to be more assertive. For this, the analysis of tasks directed by objectives, the analysis of vocabularies and properties in the criminal context will be used in order to restrict their interpretation and the evaluation of semantic weights and values of these vocabularies. Therefore, in the end, it will be possible to measure the necessary information within a risk management system.

Keywords: Situational Awareness; Ontology; Quantitative Analysis; Risks Management

1. Introdução

Consciência Situacional (*Situational Awareness* - SAW) é um conceito fundamental para auxiliar a tomada de decisão em ambientes complexos e dinâmicos em uma variedade de domínios, entre eles o de gerenciamento de riscos (BOSSÉ; ROY; WARK, 2007).

SAW é definida como a percepção dos elementos no ambiente dentro de um volume

de tempo e espaço, a compreensão do seu significado e a projeção de seu status em um futuro próximo (ENDSLEY, 1998).

Neste contexto, operadores de sistemas de gerenciamento de riscos, como os que lidam com dados criminais, e que estão constantemente sob alta pressão e expostos a uma gama de informações sensíveis, precisam manter seus níveis de SAW elevados para assim sustentar o melhor

retrato de uma situação crítica e tomar a decisão mais assertiva, evitando prejuízos à vida, ao patrimônio e também ao meio ambiente.

Por ser um estado cognitivo do operador, a SAW não garante que este tomará a melhor decisão, entretanto garante melhores subsídios para que possa melhorá-la.

Para suportar a obtenção e manutenção de SAW, foi constatado que o uso de modelos semânticos, mais especificamente ontologias, quando aplicados para suportar sistemas de gerenciamento de riscos, podem contribuir para uma melhor assertividade nas inferências úteis à tomada de decisão (MATHEUS; KOKAR; BACLAWSKI, 2003).

A tendência dos sistemas dinâmicos é utilizar ontologias ou outros modelos semânticos para que os dados sejam representados. Entretanto, esta ação gera novos desafios para quantificar os dados necessários para que assim sejam transformados em informações que apoiem à tomada de decisão dos operadores de sistemas de gerenciamento de riscos.

Fluit *et al.* (2006) apontaram a dificuldade dos operadores em fornecer termos que melhor descrevam as suas necessidades de informação. Esta dificuldade aumenta quando expressões lógicas simples são utilizadas, o que tornou o processo de quantificação mais complicado.

A quantificação é fundamental para a estruturação de um processo de análise crítica sobre os dados.

A partir dos dados oferecidos pelo sistema, deseja-se extrair tipos específicos de informações e quantificá-las para que sejam armazenadas e posteriormente analisadas por meio de métodos estatísticos, obtendo outras informações geradas a partir de dados preditivos dentro de um domínio. Sagion *et al.* (2007) afirmam que informações quantitativas (numéricas) são necessárias em sistemas dinâmicos e inteligentes.

Quando utilizada uma ontologia em um sistema inteligente, passa-se a ter mais qualificação dos dados, em vez de quantificação, tornando desejável a aplicação de métodos que supram essa necessidade.

Este trabalho tem como objetivo apresentar o desenvolvimento de um processo de quantificação de informações de

riscos, organizadas e representadas semanticamente em ontologias de domínio, mais especificamente no contexto criminal, com o principal objetivo de contribuir com a aquisição de SAW.

2. Gerenciamento de Riscos e SAW

Endsley (1995) afirma que dentro de ambientes de avaliação de riscos, é grande a necessidade de tomadas de decisões precisas e as tarefas dependem de análises contínuas e que estejam sempre atualizadas.

Um exemplo corriqueiro de SAW é um motorista de carro, o qual precisa estar ciente de todas as ações tomadas, como quando acelerar e frear, quando deve parar no semáforo e principalmente estar atento aos pedestres e carros que estejam ao seu redor, observando as condições do tráfego para que assim ele possa escolher o melhor trajeto, se possível.

A complexidade e a dinâmica de um ambiente de avaliação de riscos aumentam proporcionalmente conforme a quantidade de variáveis existentes nesse ambiente, tornando a aquisição e manutenção de SAW processos mais difíceis.

De acordo com o Modelo de Consciência Situacional proposto por Endsley (1988), existem três níveis os quais os operadores podem atingi-los conforme o seu desenvolvimento dentro do ambiente, que são influenciados direta e indiretamente por fatores internos e externos: (a) nível um: perceber a dinâmica, o estado e os atributos de elementos pertinentes ao ambiente; (b) nível dois: compreender a situação atual, criando relações entre os elementos, desenvolvendo um contexto de acordo com os objetivos e metas esperados, priorizando elementos importantes e o que estas informações combinadas representam; e (c) nível três: projetar estados futuros, tendo ciência do que os elementos são e o que significam perante a situação atual, sendo necessário ter a habilidade de prever quais serão as ações futuras.

É possível qualificar a SAW de um operador por meio de suas habilidades, preconceitos, experiências, metas e objetivos, entretanto até os mais qualificados podem tomar decisões errôneas caso sua SAW não esteja nos níveis adequados.

Em ambientes emergenciais - cujas falhas podem provocar perdas de vidas e danos ao meio ambiente e ao patrimônio público - é necessário adquirir, manter e melhorar a SAW constantemente e para isso é preciso que existam sistemas que a apoiem.

Os operadores precisam agir de forma decisiva dentro de intervalos de tempo curtos com informações muitas vezes incompletas e/ou com muitos dados dos quais é difícil extrair informações relevantes.

Os sistemas inteligentes devem fornecer informações precisas, de forma rápida e em um formato correto aos operadores. Eles devem ser elaborados de modo a auxiliarem na tomada de decisão mais assertiva pelo operador, considerando o tempo de resposta como fator limitador.

Como exemplo de ambiente emergencial, é possível utilizar o domínio militar, constatando o uso de SAW e a aplicação de seus níveis da seguinte forma: (1) percepção do operador sobre o posicionamento e condicionamento físico de sua equipe, armamento disponível e a quantidade de tropas inimigas; (2) compreensão sobre o avanço de tropas inimigas, possibilidade de deslocamento ou combate; e (3) visão de um futuro próximo em relação à tropa do inimigo, se sua atual posição representa uma ameaça ou risco iminente à equipe.

3. Ontologia como Modelo de Apoio a SAW

Ontologia é a forma como a lógica explica o significado pretendido de um vocabulário formal, ou seja, seu compromisso ontológico com uma conceitualização particular do mundo (GUARINO, 1998).

O seu uso possibilita construir relações organizadas entre diversos termos de um domínio específico, tornando possível a contextualização dos dados e auxiliando no processo de interpretação.

Kokar, Matheus e Baclawski (2009) afirmam que a diferença da obtenção e manutenção de SAW entre humanos e computadores é que no primeiro é possível mensurar, enquanto no segundo é preciso definir e implementar este processo.

As ontologias tornam-se de fato computacionais a partir do momento que elas são implementadas, algo além de conceitos, viabilizando o desenvolvimento de aplicações mais robustas e inteligentes.

Uma ontologia é valiosa somente se a maioria da comunidade envolvida em seu uso aceitar seus principais conceitos e estrutura, por isso é necessária uma definição formalizada (KOKAR; MATHEUS; BACLAWSKI, 2009).

Para utilizar SAW, especialmente para tomada de decisões, é preciso reconhecer diversas situações, avaliar o impacto sobre os objetivos, relacionar propriedades a situações particulares e comunicar as descrições destas às demais pessoas. Com isto, existem dois requisitos adicionais em relação às representações situacionais: (1) elas podem ser classificadas por tipos de situação, e (2) elas podem ser tratadas como objetos físicos ou conceituais.

Dois aspectos importantes que não podem deixar de ser considerados são os de atributos e relacionamentos, os quais devem ser associados com valores que possam mudar ao longo do tempo. Devido às várias formas que as situações podem ser abordadas, principalmente aquelas caracterizadas como emergenciais, estas devem exigir robustez e solidez da solução de SAW.

4. Metodologia

Sistemas que buscam apoiar a avaliação e análise de situações devem envolver um esforço para estimular a SAW em operadores, o mais próximo possível da realidade, uma vez que erros de SAW podem comprometer a vida ou o patrimônio. Adquirir SAW ainda é algo desafiador, frente ao grande volume, variedade e dinamicidade de dados.

Para Botega (2016), esta heterogeneidade característica dos ambientes emergenciais demandam que os dados sejam processados e representados de forma semântica, utilizando ontologias, em busca de uma informação situacional significativa.

Dentro do domínio criminal, por meio do processo de quantificação semântica, deve ser possível mensurar quais os locais de

origens mais recorrentes das denúncias, qual o horário de pico de ocorrências, a atualidade das denúncias (horário atual contra o horário da denúncia), quais as características recorrentes de vítimas, criminosos, objetos e locais.

O processo de quantificação semântica envolve o estudo de metodologias de mensuração semântica, entre eles o de Pesos e Valores Semânticos (apresentada a seguir), e a avaliação de vocabulário de integração e associação de relacionamentos semânticos, por meio da compreensão do domínio de gerenciamento de risco.

Para identificar as necessidades dos operadores foi empregada a Análise de Tarefas Dirigida por Objetivos (*Goal-Driven Task Analysis – GDTA*), uma metodologia útil para revelar as informações necessárias para que operadores humanos tomem decisões ao longo do uso de um sistema de avaliação de situações de risco, bem como as tarefas que os mesmos devem realizar. Saber as informações demandadas por cada tarefa é o primeiro passo para o processo de quantificação, pois é necessário ter conhecimento daquilo que se deseja mensurar, representando semanticamente dados que façam sentido. Parte do GDTA é apresentada no Apêndice A, a qual apresenta um de seus objetivos, que é a caracterização das quatro entidades de um crime de roubo (vítimas, criminosos, objetos e locais).

O estudo de vocabulários e propriedades utilizadas no contexto criminal é uma etapa fundamental do trabalho, pois se torna possível a descrição de uma realidade específica, restringindo sua interpretação. Estes serão criados somente após o entendimento do domínio, constituindo a semântica e conceitos. Compreendê-los como um todo proporciona um conhecimento amplo das características e propriedades de suas classes. Kokar, Matheus e Baclawski (2003) afirmam que quanto mais extenso for o vocabulário, melhor poderá ser a descrição dos relacionamentos que acontecem dentro do domínio.

A ontologia em desenvolvimento pode ser vista no Apêndice B, a qual contempla as classes principais que caracterizam um crime

de roubo e suas respectivas subclasses e relacionamentos.

Antes de quantificar os resultados por meio da representação do modelo semântico, é necessário que exista um ou mais dados, bem estruturados, cujos significados sejam relevantes.

Por fim, a quantificação em si é baseada na metodologia de Peso e Valor Semântico, proposta por Hepp *et al.* (2006), a qual leva em consideração o grau de especificidade do vocabulário e propriedades. Ela é baseada na fundamentação de que uma propriedade, quando muito utilizada, é geralmente menos específica do que uma propriedade pouco utilizada.

O primeiro passo desta metodologia é dar ao atributo do vocabulário um peso semântico, em seguida o valor semântico é calculado somando os pesos semânticos de todos os atributos. Caso uma classe não possua atributos, o valor semântico é igual zero.

Cada atributo tem realizada sua contagem de entradas entre a relação atributo e classe, portanto, segundo Hepp (2004), ele recebe um peso semântico que é igual ao seu valor recíproco de sua frequência de uso.

5. Resultados Esperados

Até o presente momento, o GDTA já foi elaborado e aplicado junto aos operadores da PMESP (Polícia Militar do Estado de São Paulo). A ontologia de domínio está em fase de desenvolvimento, embasada nas respostas obtidas do questionário do GDTA.

É esperado que com a ontologia concluída apoiando à SAW, seja possível desenvolver o processo de quantificação semântica voltado para o domínio de sua aplicação, no caso o de gerenciamento de riscos, possibilitando mensurar os relacionamentos, similaridades e atributos das classes. Uma vez mensurados, os dados podem se tornar informações valiosas que auxiliem na tomada de decisões de operadores.

4 Considerações Finais

Ao se utilizar um modelo semântico apoiado em SAW em um sistema dinâmico, é possível qualificar os dados, atribuindo um

significado ao contexto em que se encontram, porém perde-se o necessário para a quantificação dos dados: expressões lógicas e numéricas.

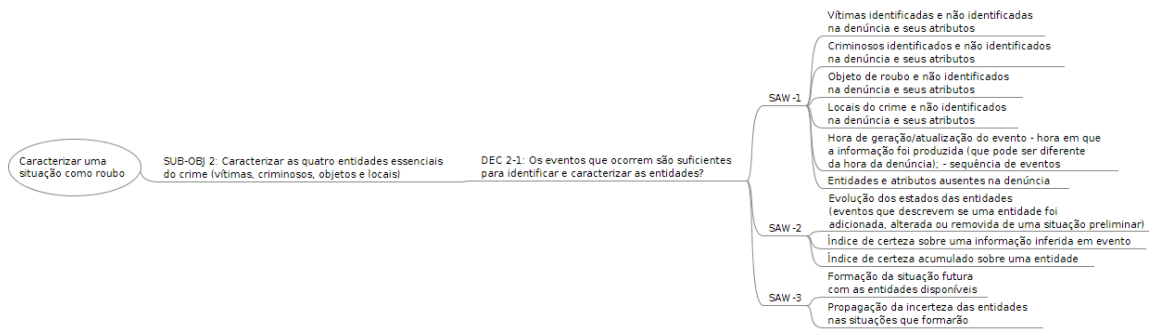
Este trabalho está em fase de desenvolvimento, principalmente a ontologia de domínio, a qual está representada no Apêndice B. Concluída, será possível representar semanticamente as informações geradas a partir de dados quantificados, servindo como suporte aos operadores de sistemas de gerenciamento de riscos com o objetivo de contribuir com a melhoria de sua SAW.

Para os próximos passos os RDFs (*Resource Description Framework* – modelo de representação de ontologia) serão gerados e analisados utilizando consultas SPARQL (*SPARQL Protocol and RDF Query Language* – linguagem de consulta semântica de dados), atribuindo os pesos semânticos e recuperando a quantificação no contexto desejado.

Referências

- BOSSÉ, É.; ROY, J.; WARK, S. **Concepts, Models, and Tools for Information Fusion**. [S.l.]: Artech House, Incorporated, 2007. (Artech House intelligence and information operations library). ISBN 9781596930810.
- BOTEGA, L. C. **Modelo de Fusão Dirigido por Humanos e Ciente de Qualidade de Informação**. 247 p. Tese (Doutorado) - UFSCar - Universidade Federal de São Carlos, 2016.
- ENDSLEY, M. R. **Toward a theory of situation awareness in dynamic systems**. Human Factors: The Journal of the Human Factors and Ergonomics Society, SAGE Publications, v. 37, n. 1, p. 32–64, 1995.
- ENDSLEY, M. R. **Design and evaluation for situation awareness enhancement**. In: SAGE PUBLICATIONS. Proceedings of the human factors and ergonomics society annual meeting. [S.l.], 1988. v. 32, n. 2, p. 97–101.
- FLUIT, Christiaan; SABOU, Marta; VAN HARMELEN, Frank. **Ontology-based information visualization: toward semantic web applications**. In: Visualizing the semantic web. Springer London, 2006. p. 45-58.
- GUARINO, N. **Formal ontology and information systems**. In: Proceedings of FOIS. [S.l.: s.n.], 1998. v. 98, n. 1998, p. 81–97.
- HEPP, Martin. **Measuring the Quality of Descriptive Languages for Products and Services**. In: MKWI (E-Business). 2004. p. 157-168.
- HEPP, Martin; LEUKEL, Joerg; SCHMITZ, Volker. **A quantitative analysis of product categorization standards: content, coverage, and maintenance of eCI@ ss, UNSPSC, eOTD, and the RosettaNet Technical Dictionary**. Knowledge and Information Systems, v. 13, n. 1, p. 77-114, 2007.
- KOKAR, Mieczyslaw M.; MATHEUS, Christopher J.; BACLAWSKI, Kenneth. **Ontology-based situation awareness**. Information fusion, v. 10, n. 1, p. 83-98, 2009.
- MATHEUS, C. J.; KOKAR, M. M.; BACLAWSKI, K. **A core ontology for situation awareness**. In: Proceedings of the Sixth International Conference on Information Fusion. [S.l.: s.n.], 2003. v. 1, p. 545–552.

Apêndice A – Análise de Tarefas Dirigida Por Objetivos (Goal-Driven Task Analysis – GDTA)



Apêndice B – Ontologia de Domínio Desenvolvida

