

APLICAÇÃO DE REGRAS DE ASSOCIAÇÃO PARA IDENTIFICAÇÃO DE PADRÕES DE OCORRÊNCIA EM ACIDENTES DO TRABALHO

Rodrigo Frank de Souza Gomes (Universidade do Vale do Rio dos Sinos)

Leandro Gauss (Universidade do Vale do Rio dos Sinos)



Definição do Problema – Apesar de maior atenção gerencial e dos esforços significativos para a prevenção de acidentes do trabalho, a ocorrência de lesões parcialmente incapacitantes ainda é um desafio para a organização.

Análise do Problema – Uma das medidas tipicamente utilizadas para redução de acidentes do trabalho é a análise das causas raízes. Este tipo de análise estimula ciclos de melhoria contínua, em que, a partir dos aprendizados gerados no processo de investigação, medidas são estabelecidas para a eliminação ou neutralização dos fatores determinantes que foram identificados na análise. Diversos métodos estão disponíveis para esse objetivo (ex: Ishikawa, 8D). Entretanto, os dados históricos da organização apontam que esse tipo de análise é necessária, mas não suficiente. Isto porque fatores potencialmente ocultos (ou o efeito combinado dos fatores mapeados) são difíceis de serem identificados por meio de técnicas mais tradicionais. Neste contexto, o uso de técnicas complementares mais robustas foram consideradas para a solução do problema.

Solução do Problema – No intuito de identificar elementos contributivos para ocorrência de acidentes, e que até então não haviam sido identificados, foi utilizada a técnica de regras de associação por meio do algoritmo Apriori proposto por Agrawal et al. (1993; 1994). Uma regra de associação é representada por um padrão do tipo $R_i: X \rightarrow Y$, onde X representa os fatores antecedentes e Y o fator consequente, neste caso o acidente do trabalho. A solução foi implementada inicialmente por meio do uso do software RStudio para análise dos resultados preliminares e validação do modelo. Posteriormente, uma arquitetura funcional foi desenvolvida para integrar a solução proposta ao sistema de gestão da existente organização.

Resultados – A partir de padrões de coocorrência revelados, foi possível identificar a associação de fatores críticos com a ocorrência de acidentes. Os resultados confirmaram associações já esperadas e também revelaram aspectos contraintuitivos até então não

considerados, como por exemplo a ocorrência significativa de acidentes envolvendo profissionais mais experientes e utilizando Equipamentos de Proteção Individual (EPIs). Outro exemplo contraintuitivo revelado é a ocorrência de acidentes com empregados treinados nos processos de prevenção de acidentes da organização, sugerindo a necessidade de revisão dos métodos aplicados em treinamentos internos. Os resultados também confirmam a relevância da qualificação técnica para a prevenção de acidentes e de fatores comportamentais a serem trabalhados, como a minimização de riscos e a falta de concentração.

Avaliação e Lições Aprendidas – A partir dos resultados encontrados, algumas lições podem ser destacadas, como por exemplo: i. inspeções técnicas no shop-floor devem ampliar o campo de análise, não se restringindo a verificação de fatores básicos como o uso de EPIs; ii. o processo de onboarding técnico deve intensificar o módulo de prevenção de acidentes, reavaliando seu conteúdo e técnicas de ensino-aprendizagem; iii. líderes organizacionais devem considerar aspectos contraintuitivos na análise de acidentes; iv. técnicas convencionais para identificação de causas raízes são necessárias, mas não suficientes para revelar fatores ocultos que individual ou combinadamente, contribuem para a ocorrência de acidentes.

Organização/Empresa – TK Elevator é uma das empresas líderes mundiais na indústria de fabricação, instalação e manutenção de elevadores, escadas rolantes e equipamentos de mobilidade urbana. O objeto do caso utilizado neste artigo foi a Business Unit Latin America, presente em 12 países.

CNAE (Classificação Nacional de Atividade Econômica): 28.22-4

Palavras-chave: Acidente do trabalho. Prevenção de acidentes. Regras de associação.

1. Introdução

A gestão de saúde e segurança do trabalho (SST) tornou-se uma questão global e soluções para melhorar seu desempenho têm sido requisitadas em ambientes de trabalho modernos (WANG *et al.*, 2020). Isto ocorre porque apesar da maior atenção gerencial e da melhoria progressiva das práticas de SST ao longo de décadas, muitos desafios ainda permanecem sem solução, em particular os acidentes e doenças relacionadas ao trabalho (NICOLAIDOU *et al.*, 2021).

Segundo a Organização Internacional do Trabalho (ILO, 2020), mais de 2,8 milhões de mortes e aproximadamente 376,8 milhões de acidentes não-fatais ocorrem todos os anos. Esse cenário incômodo sugere que as diretrizes propostas na agenda internacional (ILO, 1999; UN, 2015) não tem sido suficientes para avanços significativos na promoção da saúde e a segurança de trabalhadores, e que as práticas de SST adotadas nas organizações não tem sido efetivas na prevenção de acidentes.

Entre as medidas tipicamente utilizadas pelas organizações para redução de acidentes, se encontra a análise de causas raízes (ISO, 2018). Este tipo de análise estimula ciclos de melhoria contínua, em que, a partir dos aprendizados gerados no processo de investigação, medidas são estabelecidas para a eliminação ou neutralização dos fatores determinantes que foram identificados na análise.

Diversos métodos estão disponíveis para esse objetivo (ex: Ishikawa, 8D). Entretanto, a análise isolada de acidentes não permite a identificação de padrões de ocorrência estatisticamente relevantes. Isto ocorre porque tais métodos não consideram o conjunto de dados históricos e, portanto, a coocorrência de fatores antecedentes e consequentes são difíceis de serem identificados.

Esse estudo considera o uso de técnicas mais robustas com o objetivo de identificar padrões de ocorrência válidos que possibilitem ações gerenciais para a prevenção de acidentes. Assim, a técnica de mineração de dados por regras de associação proposta por Agrawal *et al.* (1993; 1994) é aplicada em um caso prático onde são revelados fatores antecedentes considerados prováveis de co-ocorrerem com acidentes (fator consequente).

O artigo está estruturado como segue: a seção 2 aborda os aspectos teóricos da técnica empregada. Na seção 3, o método de implementação é apresentado por meio do detalhamento da arquitetura funcional utilizada. Na seção 4 são apresentados e discutidos os resultados. Finalmente, as conclusões e limitações do estudo constam na seção 5.

2. Mineração de dados por Regras de Associação

A mineração de dados por regras de associação tem por objetivo identificar padrões significativos de coocorrência entre variáveis que compõem um objeto de estudo, como por exemplo, a associação entre fatores individuais e a ocorrência de acidentes (BARKER, 2021). A técnica foi inicialmente proposta por Agrawal *et al.* (1993), quando formularam um algoritmo de rápido processamento para identificar coocorrência no consumo de produtos em uma organização de varejo. A partir do registro de transações comerciais, o algoritmo foi capaz de identificar padrões de ocorrência do tipo: “quando o consumidor compra os produtos *A* e *B* (*fatores antecedentes*), também compra o produto *C* (*fator consequente*)”. Posteriormente, considerando seu elevado potencial de aplicação em diversas outras áreas, o algoritmo então chamado *Apriori* foi modificado para que fosse possível identificar a regras de associação com múltiplos fatores antecedentes e múltiplos fatores consequentes (AGRAWAL; SRIKANT, 1994).

Em um conjunto de dados D que contém transações do tipo $T = \{t_n | n = 1, \dots, N\}$, uma regra de associação é expressa na forma $R_m: \{X\} \rightarrow \{Y\}$, em que $X = \{x_i | i = 1, \dots, I\}$ e $Y = \{y_k | k = 1, \dots, K\}$ representam fatores antecedentes e consequentes, respectivamente.

A significância de uma regra de associação em D é definida por meio de parâmetros estatísticos chamados de suporte (*supp*) e confiança (*conf*). Estes parâmetros de referência são utilizados para capturar um certo nível de relacionamento entre itens presentes em um conjunto de dados (LIAO; PERNG, 2008).

O suporte é definido como a probabilidade de X e Y coexistirem no conjunto de dados, ou ainda, a fração da quantidade de transações no conjunto de dados que contém todos os itens em uma regra específica, i.e., $supp(X \rightarrow Y) = \sigma(X \cap Y)/N$. Para regras de associação, um limite mínimo de suporte (*minsupp*) é usado para selecionar combinações de fatores frequentes (e esperançosamente importantes) (HAHSLER; HORNIK, 2008).

Já a confiança (*conf*) representa uma medida de confiabilidade da regra, ou seja, quando maior a confiança de $X \rightarrow Y$, maior a probabilidade de Y estar presente nas transações que contém X , tal que $conf(X \rightarrow Y) = \sigma(X \cap Y)/\sigma(X)$.

Uma regra de associação $R_m: \{X\} \rightarrow \{Y\}$ é considerada relevante se pelo menos satisfizer os limites mínimos de suporte ($supp \geq minsupp$) e confiança ($conf \geq minconf$) definidos no modelo (LIAO; PERNG, 2008). Tais parâmetros de referência podem ser encontrados em estudos disponíveis na literatura ou serem definidos sob o interesse do analista (ALVES, 2020;

BARALIS; PSAILA, 1997; ISA *et al.*, 2018; KOURIS; MAKRIS; TSAKALIDIS, 2005; ZHANG; ZHANG, 2002).

Entretanto, a depender do conjunto de dados, o volume de regras de associação geradas pode ser muito significativo apesar de atenderem aos critérios mínimos estabelecidos pelos parâmetros *supp* e *conf* utilizados. Uma forma de qualificar a análise é considerar uma medida de interesse chamada lift (*l*). Essa medida compara a frequência em que *X* e *Y* ocorrem juntos com a frequência esperada se fossem estatisticamente independentes: $lift = supp(X \rightarrow Y) / supp(X) \cdot supp(Y)$ (HAHSLER, 2015). Quando o lift de uma regra apresenta valores superiores a 1, há indicação de que os fatores estão positivamente relacionados (quanto maior o lift maior a força da regra). Para valores de lift menores que 1, os fatores possuem uma fraca relação entre si. Um valor de lift de 1 indica independência entre *X* e *Y*.

Assim, as regras de associação são representadas conforme elementos presentes na Tabela 1.

Tabela 1 – Estrutura de Regras de Associação

Regra de Associação	Fator antecedente (<i>X</i>)		Fator consequente (<i>Y</i>)	<i>supp</i>	<i>conf</i>	<i>lift</i>
R1	{A, B}	=>	{C}	[0 – 1]	[0 – 1]	[0 - ∞]
R2	{A, D, E}	=>	{F, G}	[0 – 1]	[0 – 1]	[0 - ∞]

Fonte: Adaptado de Gomes *et al.* (2022)

Os resultados gerados com o uso desta técnica para identificação de padrões ocultos de coocorrência de fatores tem chamado a atenção de acadêmicos e profissionais de SST. Isto se evidencia pelo crescente volume de estudos com aplicações de regras de associação na prevenção de acidentes, em particular na indústria da construção, ferroviária e no controle de estradas e rodagem (LIAO; PERNG, 2008; MIRABADI; SHARIFIAN, 2010; MONTELLA, 2011). Entretanto, a literatura ainda é escassa em aplicações que integram essa técnica na arquitetura do sistema de gestão de saúde e segurança organizacional. Na próxima seção será apresentado o método de aplicação no caso em estudo.

3. Etapas de implementação da solução proposta

Este estudo empregou o método *Knowledge Discovery in Databases* (KDD) proposto por Fayyad (1996). KDD significa a descoberta de conhecimento em bancos de dados. Trata-se, portanto, de um processo não trivial de identificar padrões válidos, novos, potencialmente úteis e, em última análise, compreensíveis em dados (FAYYAD *et al.*, 1996).

Na etapa inicial, os dados utilizados na pesquisa foram coletados a partir de relatórios de investigação de acidentes ocorridos no período entre outubro de 2020 e fevereiro de 2022. No total, 74 relatórios de investigação foram mapeados nas operações da TK Elevator na América Latina, que inclui 12 países, dezenas de unidades operacionais e uma fábrica.

Na etapa seguinte, os dados foram compilados em uma tabela composta por fatores e atributos considerados para a compreensão de causas raízes, tais como caracterização do local do acidente, condições de trabalho, fatores individuais e organizacionais. A lista de fatores e atributos está apresentada na Tabela 2 e foi definida com base em estudos anteriores (CHENG; LIN; LEU, 2010; VERMA *et al.*, 2014) e complementada por especialistas da organização.

Tabela 2 – Fatores e atributos

Seq.	Fator	Atributo
1	Tipo de acidente	Com perda de tempo; sem perda de tempo; fatal
2	Área	Manutenção; obras novas; modernização
3	Tipo de contrato	Empregado; prestador de serviços
4	Escolaridade	Fundamental; médio; superior
5	Formação técnica	Sem formação; mecânica; elétrica; outras
6	Função	Técnico de manutenção; montador; etc
7	Anos de experiência	<1 ano; entre 1 e 5 anos; mais que 5 anos; mais que 10 anos
8	Equipamento	Elevador; escada rolante; acessibilidade; outros
9	Dia da semana/turno	segunda-feira, terça-feira, ... / manhã, tarde, noite
10	Condição da tarefa	Rotineira; não-rotineira
11	Status do projeto	On-time; atrasado
12	Localização do acidente	Casa de máquinas, caixa de corrida, poço, outros
13	Horário de trabalho no acid.	Regular; extra
14	Análise de risco (1)	APR realizada; APR não realizada
15	Análise de risco (2)	APR adequada; APR não adequada
16	Uso de ferramentas	Adequado; não adequado
17	Uso de EPIs	EPIs em uso; EPIs não-conforme
18	Procedimento padrão de trabalho	Existente; inexistente
19	Procedimento padrão de trabalho utilizado	Conforme; não-conforme
20	Treinamento para a tarefa	Sim; não
21	Histórico prévio de acidentes	Sim; não
22	Histórico prévio de advertências	Sim; não
23	Auditorias realizadas nos últimos 12m	Sim; não
24	Observação comport. nos últimos 12m	Sim; não
25	Mudança comportamental recente	Sim; não
26	Teste psicológico realizado antes do trab.	Sim; não
27	Causa primária do acidente	Falha de processo; desvio de comportamento
28	Tipo associado com causa primária	Condição inseguro; ato inseguro; ato inseguro de outros

29	Associação com regra de segurança violada	Proteção contra queda, circuitos elétricos, barricadas, outros
30	Associação com armadilha comportamental	Falta de concentração, falta de conscientização, outros

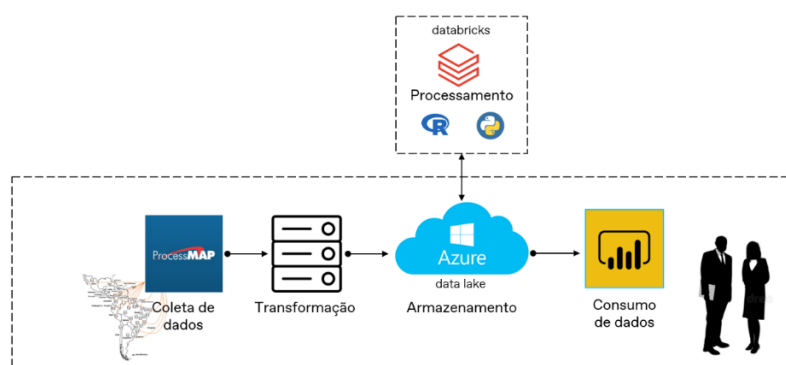
Fonte: Elaborado pelos autores

A partir dos dados estruturados (Tabela 2), o algoritmo *Apriori* foi utilizado para gerar as regras de associação mais relevantes. Nesta etapa, o *software* RStudio (2020) foi utilizado para processamento do código (ver apêndice A 1) a partir de parâmetros pré-estabelecidos de suporte ($supp = 0.2$) e confiança ($conf = 0.5$). A definição dos parâmetros levou em consideração referências na literatura e o interesse da organização em não restringir excessivamente o número de regras geradas como forma de garantir uma ampla pesquisa nos dados.

Após análise e validação do modelo por especialistas em SST da organização, a etapa seguinte foi a definição e construção de uma arquitetura funcional que realiza a integração do processo de mineração de dados proveniente do algoritmo *Apriori* com os sistemas existentes na organização (Figura 1).

Em termos simples, a arquitetura proposta possui dois fluxos. O primeiro conecta uma fonte geradora de arquivos primários (ex: dados provenientes dos usuários que reportam incidentes em todas as localidades onde a empresa possui operação na América Latina utilizando o *software* ProcessMAP), os transforma para um padrão seguro e estruturado a partir da análise e investigação, e armazena os dados no *data lake* (termo comumente utilizado para expressar o local onde dados já transformados encontram-se armazenados). A partir do *data lake*, um segundo fluxo é acionado para processamento do algoritmo *Apriori* através de uma plataforma que unifica bases de dados, permitindo análises que utilizam múltiplas linguagens computacionais (*databricks*). Ao final, os resultados gerados (regras de associação) retornam ao *data lake* e são apresentados no Power B.I. para consumo.

Figura 1 - Arquitetura funcional para processamento de Regras de Associação



Fonte: Elaborado pelos autores

A partir da construção da arquitetura funcional desenvolvida, todos os relatórios de investigação de acidentes são estruturados e adicionados em um mesmo banco de dados, que por sua vez, atualiza as regras de associação relevantes e as apresenta na forma de um *dashboard* disponibilizado para gerentes operacionais e especialistas de SST da organização. Os resultados provenientes dessa arquitetura são apresentados e discutidos na seção seguinte.

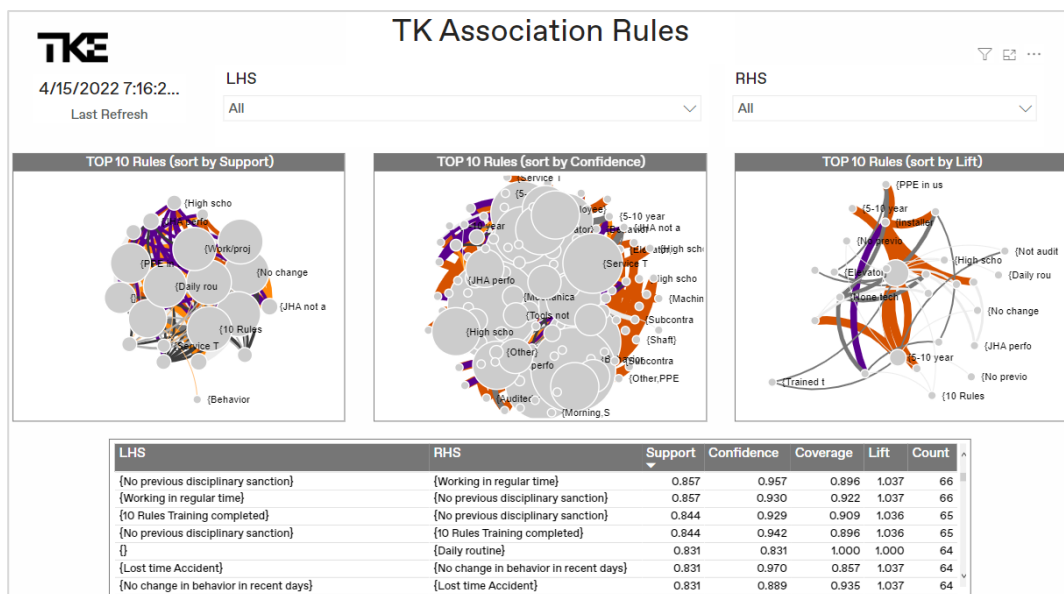
4. Resultados

4.1 Arquitetura funcional

A definição de uma arquitetura funcional integrada permitiu rápida aderência ao consumo dos dados por parte de especialistas em SST e gerentes operacionais. Algumas vantagens dessa arquitetura incluem: (i) visualização do processo de mineração de dados (Figura 2); (ii) fácil e rápido acesso as regras de associação geradas por cada um dos parâmetros selecionados (Figura 3); e (iii) possibilidade de navegar entre diferentes análises a partir da seleção de fatores consequentes de interesse (campo RHS).

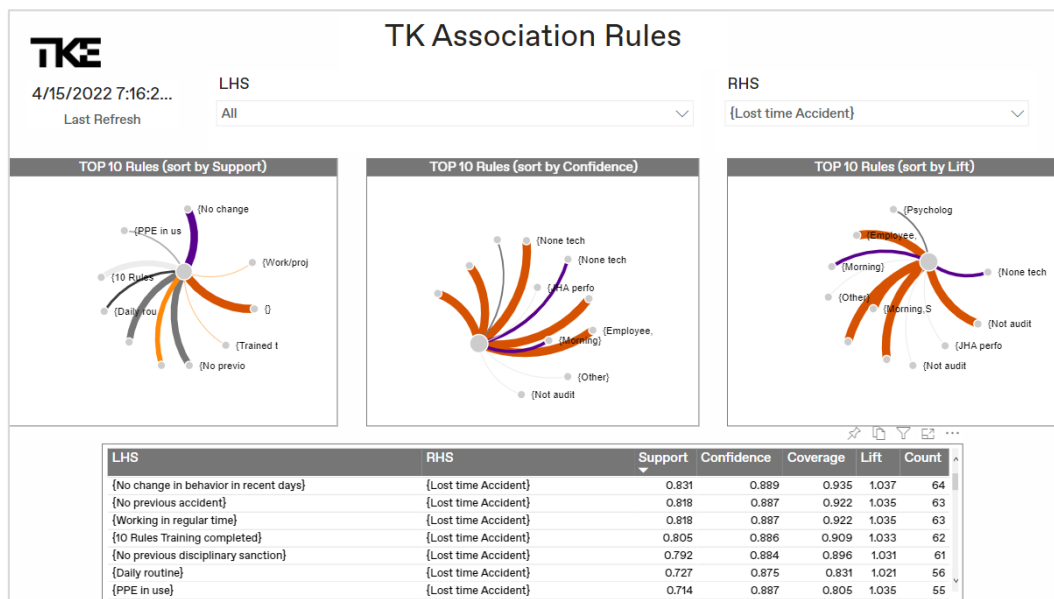
As Figuras 2 e 3 apresentam o resultado da arquitetura desenvolvida. Os atributos que compõem fatores analisados nas investigações de acidentes são processados de maneira que sejam identificadas a força das associações, expressa no *dashboard* por meio da espessura e cor da linha, assim como através do tamanho do círculo. Por exemplo, quanto maior o círculo maior a frequência da coocorrência entre 2 atributos, assim como quanto mais espessa a linha de conexão entre atributos, maior a confiança ou o lift.

Figura 2 – Mineração de dados em processamento (screenshot)



Fonte: TK Elevator

Figura 3 – TOP 10 regras de associação para RHS = Acidente com perda de tempo



Fonte: TK Elevator

4.2 Regras de associação e padrões de ocorrência

Um total de 39 regras de associação foram geradas a partir dos parâmetros de suporte e confiança estabelecidos para o modelo. Como resultado, alguns padrões de ocorrência já esperados foram confirmados, assim como outros contraintuitivos foram revelados. A seguir estão apresentadas as regras de associação cujo impacto na prevenção de acidente é considerado significativo:

- Os acidentes ocorrem de forma prevalente em horário normal de trabalho ($supp = 0,818$) e em atividades de rotinas ($supp = 0,727$);
- ~70% dos acidentes ocorrem com colaboradores não auditados nos últimos 3 meses ($supp = 0,662$);
- Uma quantidade considerável dos acidentados não tinha histórico prévio de acidentes no trabalho ($supp = 0,818$), nem mesmo de advertências disciplinares anteriores ao acidente ($supp = 0,79$);
- Os acidentados estavam treinados nas regras de segurança estabelecidas pela organização ($supp = 0,805$);
- Em mais de 60% das vezes a análise preliminar de risco não foi conduzida adequadamente ($supp = 0,623$);
- A ausência de formação técnica está associada com acidentes em elevado nível de confiança ($supp = 0,390$; $conf = 0,968$);

- g) Acidentes envolvendo trabalhadores com mais de 10 anos de experiência e utilizando EPIs compõem um padrão de ocorrência relevante ($conf = 0,944$; $l = 1,102$);
- h) A coocorrência entre uma análise de riscos inadequada conduzida por um profissional subcontratado representa uma associação relevante para acidentes ($conf = 0,944$; $l = 1,102$);
- i) A condição de trabalhador não auditado nos últimos 3 meses ainda que com ferramentas de trabalho adequadas é uma combinação potencial para ocorrência de acidentes ($conf = 0,950$; $l = 1,108$);
- j) A condição de trabalhador não auditado nos últimos 3 meses (sendo este um subcontratado) possui elevada força de associação com acidentes ($conf = 0,995$; $l = 1,114$);
- k) Coocorrência entre a ausência de formação técnica do trabalhador e ausência de teste psicológico representa um risco para a organização ($conf = 1,000$; $l = 1,167$);
- l) Acidentes ocorrem mesmo após a realização de auditorias recentes de observação comportamental ($conf = 1,000$; $l = 1,167$).

Todas as regras de associação selecionadas na análise foram consideradas como padrões de ocorrência válidos para direcionar planos mais efetivos na redução de acidentes.

Adicionalmente, as regras revelaram alguns padrões contraintuitivos de acidentes que ocorrem com trabalhadores treinados ou ainda utilizando EPIs. A seção seguinte discute os resultados encontrados a partir da perspectiva de que a prevenção de acidentes pode ser alcançada por meio da interrupção de um padrão de ocorrência que resulta em um fator indesejado.

4.3 Prevenção de acidentes a partir de padrões de ocorrência

Acidentes do trabalho estão usualmente associados com a coocorrência de múltiplos fatores, entre os quais os desvios de procedimentos, análise de risco inadequada e a falta de uso de equipamentos de proteção individual.

Esses fatores contributivos para a ocorrência de acidentes são usualmente identificados nas organizações por meio da análise das causas raízes. Entretanto, apesar da importância para o processo de prevenção, a análise isolada de acidentes não permite a identificação de padrões de ocorrência, uma vez que não considera o conjunto de dados históricos.

Neste contexto, a combinação dos fatores (chamados antecedentes) que coocorrem com um determinado fator consequente indesejado (ex: acidente) pode ser identificada por meio de técnicas mais avançadas, como a mineração de dados por regras de associação.

Neste estudo, as regras de associação revelaram padrões de ocorrência de fatores que “resultaram” em um acidente do trabalho (ex: desvio de processo combinado com a falta de uso de EPIs). Importante destacar que não devemos assumir a existência de causalidade (se “A” ocorre então “B” ocorre”). Em vez disso, desejamos destacar a existência de uma forte associação entre atributos que coexistem nos registros dos dados analisados (se “A” ocorre então é provável que “B” também ocorre).

Conforme anteriormente destacado, o conjunto de regras de associação resultantes da análise realizada nos 74 acidentes investigados confirmaram condições já esperadas pelos especialistas em SST da organização, assim como revelam outras condições chamadas contraintuitivas.

Por um lado, alguns fatores esperados antes do estudo foram confirmados por meio das regras de associação, como por exemplo a forte associação entre acidentes de trabalho e trabalhadores sem qualificação técnica (item f), ou ainda a coocorrência de acidentes com análises inadequadas de riscos (item e).

Por outro lado, alguns padrões contraintuitivos despertaram a atenção dos especialistas e gestores da organização. Esses padrões até então ocultos trazem benefícios significativos para o processo de prevenção de acidentes. Como exemplo, a ocorrência de acidentes com trabalhadores sem histórico prévio de acidentes ou de sanções disciplinares, demonstrando que acidentes não ocorrem com trabalhadores de perfil caracterizado por indisciplina recorrente, mas sim, com aqueles até então isentos de histórico negativo (item c).

Outro padrão contraintuitivo é a ocorrência de acidentes com trabalhadores utilizando EPIs. Isso demonstra que o simples fato de utilizar EPIs não assegura um trabalho livre de lesões. O estudo chama a atenção para aspectos como o excesso de confiança, tipicamente relacionado com experiência e uso de EPIs (item g).

Outras regras revelaram, por exemplo, a ineficácia dos treinamentos de regras de segurança, a importância do teste psicológico em trabalhadores subcontratados, a relevância da análise de riscos, e a importância da frequência de auditorias na prevenção de acidentes.

Todos os padrões de ocorrência mencionados na seção anterior são estatisticamente relevantes, e possibilitam que ações gerenciais sejam estabelecidas para a prevenção de acidentes.

Este processo pode ser representado através da eliminação ou neutralização da coocorrência dos fatores antecedentes. Por exemplo, uma vez identificado a forte associação entre acidentes e a ausência de auditorias nos últimos 3 meses em obras com a presença de subcontratados, a organização deve estabelecer processos que asseguram auditorias mais frequentes, de forma a eliminar ou neutralizar esse padrão de ocorrência cujo resultado é indesejado.

Neste contexto, uma vez considerado o acidente do trabalho como fator consequente indesejado, todos os esforços devem ser direcionados a evitar a ocorrência dos fatores antecedentes identificados nas regras de associação relevantes.

5. Conclusão

Este estudo utilizou a técnica de mineração de dados por regras de associação para identificação de padrões de ocorrência potencialmente úteis para prevenção de acidentes do trabalho.

O modelo foi aplicado em um conjunto de dados composto por relatórios de investigação de 74 acidentes ocorridos entre outubro de 2020 e fevereiro de 2022 na organização objeto de estudo. Como resultado, um total de 39 regras de associação foram geradas a partir dos parâmetros estabelecidos para o modelo, onde padrões de ocorrência já esperados foram confirmados, e outros contraintuitivos foram revelados.

A identificação de padrões de ocorrência estatisticamente relevantes possibilita que ações gerenciais sejam estabelecidas para a prevenção de acidentes. Esse processo pode ser representado através da eliminação ou neutralização da coocorrência dos fatores antecedentes, quando o fator consequente é o acidente do trabalho.

O estudo possui contribuições teóricas e práticas. Primeiro, uma arquitetura funcional é apresentada como mecanismo de integração do algoritmo *Apriori* em sistemas organizacionais existentes, fechando uma lacuna existente na literatura. Segundo, o estudo direciona na prática como utilizar padrões de ocorrência revelados por regras de associação na prevenção de acidentes, ainda que estudos complementares sejam recomendados.

REFERÊNCIAS

- AGRAWAL, R.; IMIELINSKI, T.; SWAMI, A. Mining Association in Large Databases. **Proceedings of the 1993 ACM SIGMOD international conference on Management of data**, [s. l.], p. 207–216, 1993.
- AGRAWAL, R.; SRIKANT, R. Fast Algorithms for Mining Association. **Proceedings of the 20th VLDB Conference**, [s. l.], p. 487–499, 1994.
- ALVES, Francielle Laura Dutra. **Data mining como suporte a tomada de decisão na priorização de estoques**. 2020. 87 f. [s. l.], 2020.
- BARALIS, Elena; PSAILA, Giuseppe. Designing Templates for Mining Association. **Journal of Intelligent Information Systems**, [s. l.], v. 32, n. 9, p. 7–32, 1997.
- BARKER, Thomas T. Finding Pluto: An Analytics-Based Approach to Safety Data Ecosystems. **Safety and Health at Work**, [s. l.], v. 12, n. 1, p. 1–9, 2021. Disponível em: <https://doi.org/10.1016/j.shaw.2020.09.010>.
- CHENG, C.-W.; LIN, C.-C.; LEU, S.-S. Use of association rules to explore cause-effect relationships in occupational accidents in the Taiwan construction industry. **Safety Science**, [s. l.], v. 48, n. 4, p. 436–444, 2010.

- FAYYAD, Usama M. A concerted effort to develop knowledge. **IEEE Expert: Intelligent Systems and Their Applications**, [s. l.], v. 11, n. 5, p. 20–25, 1996.
- FAYYAD, U. *et al.* U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, From Data Mining to Knowledge Discovery: An Overview. *In: U. FAYYAD et al. (org.). Advances in Knowledge Discovery and Data Mining.* Cambridge, MA, USA: MIT Press, 1996. p. 1–36.
- GOMES, Rodrigo F S *et al.* Safety at Work: a Complex or an Exceedingly Simple Matter?. **Reliability: Theory and Applications**, [s. l.], v. 17, n. 67, p. 267–287, 2022.
- HAHSLER, Michael. A Probabilistic Comparison of Commonly Used Interest Measures for Association Rules. [s. l.], 2015. Disponível em: <https://mhahsler.github.io/arules/docs/measures>.
- HAHSLER, Michael; HORNIK, Kurt. New probabilistic interest measures for association rules. **Intelligent Data Analysis**, [s. l.], v. 11, n. 5, p. 437–455, 2008.
- ILO. **International Labour Organization Web Page**. [S. l.], 2020. Disponível em: <https://www.ilo.org/global/topics/safety-and-health-at-work/lang--en/index.htm>. Acesso em: 15 jul. 2020.
- ILO. Report of the Director-General: Decent Work. *In: , 1999, Geneva, Switzerland. 87th Session, International Labour Conference.* Geneva, Switzerland: [s. n.], 1999. Disponível em: <https://www.ilo.org/public/english/standards/reim/ilc/ilc87/rep-i.htm>. Acesso em: 2 mar. 2022.
- ISA, N *et al.* Market Basket Analysis of Customer Buying Patterns at Corm Café. **International Journal of Engineering & Technology**, [s. l.], v. 7, n. October, p. 119–123, 2018.
- ISO. **ISO 45001. Occupational health and safety management systems – Requirements with guidance for use**. [S. l.]: International Organization for Standardization, 2018. Disponível em: <https://www.iso.org/standard/63787.html>.
- KOURIS, Ioannis N.; MAKRIS, Christos H.; TSAKALIDIS, Athanasios K. Using Information Retrieval techniques for supporting data mining. **Data & Knowledge Engineering**, [s. l.], v. 52, n. 3, p. 353–383, 2005.
- LIAO, Chia Wen; PERNG, Yeng Horng. Data mining for occupational injuries in the Taiwan construction industry. **Safety Science**, [s. l.], v. 46, n. 7, p. 1091–1102, 2008.
- MIRABADI, A; SHARIFIAN, S. Application of association rules in Iranian Railways (RAI) accident data analysis. **Safety Science**, [s. l.], v. 48, n. 10, p. 1427–1435, 2010. Disponível em: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-77956228177&doi=10.1016%2Fj.ssci.2010.06.006&partnerID=40&md5=b408a27f01576dbae84259c788a5fbf>.
- MONTELLA, A. Identifying crash contributory factors at urban roundabouts and using association rules to explore their relationships to different crash types. **Accident Analysis and Prevention**, [s. l.], v. 43, n. 4, p. 1451–1463, 2011. Disponível em: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-79955577324&doi=10.1016%2Fj.aap.2011.02.023&partnerID=40&md5=c02164f58988ec5949dbf636ee042ed>.
- NICOLAIDOU, Olga *et al.* The use of weak signals in occupational safety and health: An investigation. **Safety Science**, [s. l.], v. 139, n. January, 2021.
- RSTUDIO. **RStudio: Integrated Development for R. RStudio, PBC, Boston, MA**. [S. l.: s. n.], 2020. Disponível em: <http://www.rstudio.com/>.
- UN. **Resolution adopted by the General Assembly on 25 September 2015**. [S. l.: s. n.], 2015.
- VERMA, A *et al.* Identifying patterns of safety related incidents in a steel plant using association rule mining of incident investigation reports. **Safety Science**, [s. l.], v. 70, p. 89–98, 2014.

WANG, Y *et al.* A systematic review on the research progress and evolving trends of occupational health and safety management: A bibliometric analysis of mapping knowledge domains. **Frontiers in Public Health**, [s. l.], v. 8, 2020.

ZHANG, Shichao; ZHANG, Chengqi. **Association Rule Mining - Models and Algorithms. Lecture Notes in Artificial Intelligence**. Berlin: Springer, 2002. *E-book*. Disponível em: <https://www.springer.com/gp/book/9783540435334>. Acesso em: 11 out. 2020.

APÊNDICE

A 1 – Script R for association rules

R Studio v. 4.0.5

```
# Require packages
if(!require(readxl)) install.packages("readxl")
if(!require(arules)) install.packages("arules")
if(!require(arulesViz)) install.packages("arulesViz")
if(!require(tidyr)) install.packages("tidyr")

# Load packages
library(readxl); library(arules), library(arulesViz), library(tidyr)

# Load dataset
data <- read_excel("Lost-time accidents Report.xlsx", sheet='DATA')
View(data)

# Adjust dataset
data_aj <- dados [, c(-2,-3,-4,-5,-6,-7,-8,-9)]
View(data_aj)

# Convert dataset into file .csv
write.csv(dados_aj,"AR.csv", quote=FALSE, row.names=FALSE)

# Convert dataset into transaction format
tr <- read.transactions('AR.csv', format = 'basket', sep=',')
tr
summary(tr)

# Create association rules
rules = apriori(tr, parameter=list(suppor = 0.5, conf = 0.8, minlen = 1, maxlen = 3))
rules
inspect(head(rules))

# Remove redundant rules
rules = rules[!is.redundant(rules)]
rules
inspect(rules)
result = inspect(rules)

# Print association rules
write.csv2(result, "Association rules.csv")
```