

DETECÇÃO DE ANOMALIAS NOS PAGAMENTOS DO SISTEMA ÚNICO DE SAÚDE

ANOMALY DETECTION IN THE SISTEMA ÚNICO DE SAÚDE PAYMENTS

DETECCIÓN DE ANOMALÍAS EN LOS PAGOS DEL SISTEMA ÚNICO DE SAÚDE

Renato M. Assunção¹, Osvaldo S. F. Carvalho², Marcos O. Prates³ e Marcelo Almeida Campos⁴

¹Professor Titular, Departamento de Ciência da Computação, Universidade Federal de Minas Gerais

²Professor Associado, Departamento de Ciência da Computação, Universidade Federal de Minas Gerais

³Professor Associado, Departamento de Estatística, Universidade Federal de Minas Gerais

⁴Médico, Secretaria Municipal de Saúde de Belo Horizonte

RESUMO: Objetivos: Este artigo descreve o INFOSAS, um sistema moderno, interativo e automático de detecção de discrepâncias no sistema de pagamento aos prestadores de serviços ao Sistema Único de Saúde (SUS) para posterior auditoria e verificação. **Métodos:** Algoritmos estatísticos de mineração de dados são aplicados aos dados do SUS cobrindo 269 grupos de procedimentos médicos em 5570 municípios e mais de 23 mil prestadores de saúde, num total de mais de 1.5 milhões de séries temporais. **Resultados:** Encontramos 6811 prestadores com valores considerados excedentes e discrepantes. Este grupo de prestadores é bastante desigual. O valor excedente concentrado nos 100 prestadores mais críticos é de 210 milhões de reais, ou 51% do total excedente estimado. **Conclusão:** O sistema INFOSAS pode ser utilizado no processo de indicação de casos para auditoria, melhorando a sua qualidade e reduzindo a frequência de auditorias desnecessárias.

Descritores: Mineração de dados; Auditoria Financeira.

ABSTRACT: Aims: This paper describes INFOSAS, a modern, interactive and automatic outlier detection in the payment system to the Sistema Único de Saúde (SUS) services providers for subsequent audit and verification. **Methods:** We apply statistical data mining algorithms to SUS data covering 269 groups of medical procedures in 5570 municipalities and more than 23,000 health care providers, summing up more than 1.5 million time series. **Results:** We found 6811 providers with amounts considered excessive and discrepant. This group of providers is quite uneven. The surplus value concentrated in the 100 most critical providers is 210 million of reais, or 51% of the total surplus estimated. **Conclusion:** The INFOSAS system can be used to point out to cases for auditing process, improving their quality and reducing the frequency of unnecessary audits.

Keywords: Data Mining; Financial Audit.

Introdução

A grande dimensão do setor de saúde e o enorme volume de dinheiro envolvido fazem dele um atraente alvo para fraudes em todo o mundo. Nos Estados Unidos, a National Health Care Anti-Fraud Association (NHCAA) estima que mais de \$60 bilhões são perdidos anualmente por fraude¹. O Brasil

gerencia o maior sistema de saúde pública do mundo, o Sistema Único de Saúde (SUS). Ele realiza mais de 3.5 bilhões de procedimentos ambulatoriais e mais de 12 milhões de internações hospitalares por ano, gastando quase 26 bilhões de reais com estes atendimentos. Mais de 220 mil estabelecimentos prestam serviços ao SUS, entre hospitais, clínicas e laboratórios. É óbvia a necessidade de um sistema de regulação e controle efetivo para cuidar de um sistema deste porte. Este sistema existe e funciona nas várias esferas em que o SUS se organiza, municipal, estadual e federal.

Em 2013, procurando incorporar ferramentas de mineração de dados neste sistema de controle, o Ministério da Saúde contratou uma equipe do Departamento de Ciência da Computação (DCC) da Universidade Federal de Minas Gerais (UFMG), sob a coordenação e liderança do segundo autor deste artigo, para desenvolver um sistema moderno, interativo e automático de detecção de anomalias para posterior auditoria e verificação.

A produção do SUS é mensalmente registrada nas bases Sistema de Informação Ambulatorial (SIA)² e Sistema de Informação Hospitalar (SIH)³, com informações sobre atendimentos ambulatoriais e hospitalares, respectivamente. O Ministério da Saúde mantém também uma base de dados com cadastro de estabelecimentos, o CNES⁴. O INFOSAS utiliza também dados populacionais do IBGE.

A Figura 1 mostra a arquitetura do sistema INFOSAS. Mensalmente, os dados de produção e de população alimentam uma tabela com fatos básicos, que por sua vez alimenta algoritmos de mineração. Usando esses algoritmos, o INFOSAS produz o que chamamos de *folhas de fatos*. Uma folha de fatos mostra tabelas e gráficos que relacionam um período de análise, um estabelecimento e um *alvo de mineração*.

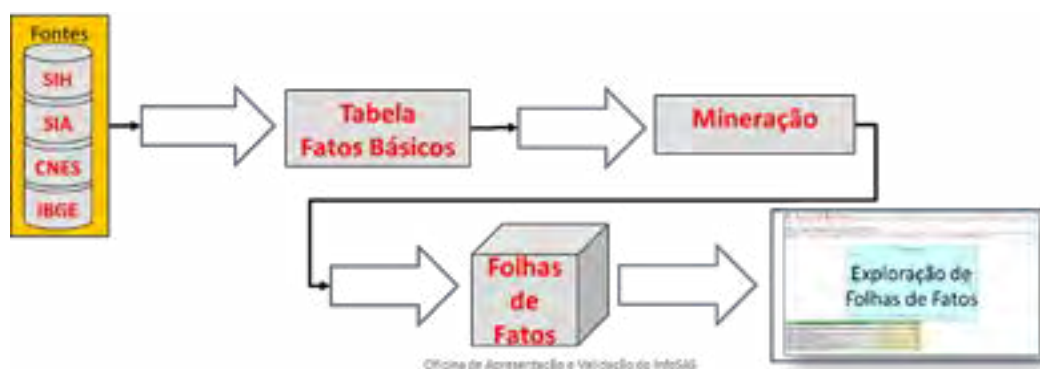


Figura 1: Arquitetura do INFOSAS

Alvos de mineração são filtros sobre os dados de atendimento do SUS, definidos tipicamente por conjuntos de procedimentos da tabela do SUS⁵. Um exemplo de alvo de mineração é o MC Cirurgia Parede Torácica (041204), definido como “*forma de organização de internação é igual a parede torácica e a complexidade é igual a média*”. Outro exemplo de alvo é Diagnóstico por Ressonância Magnética (0207), definido como “*sub-grupo ambulatorial igual a diagnóstico por ressonância magnética com o município do paciente preenchido*”.

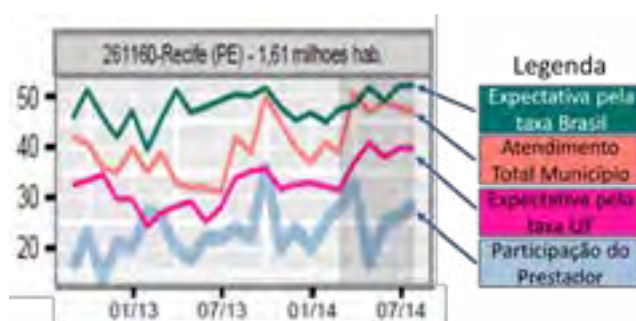


Figura 2: Dados de cobertura populacional para o alvo SIH UTI - AC - Cirurgia cardiovascular - Cardiologia intervencionista (040603)

A Figura 2 mostra um exemplo de gráfico presente em folhas de fatos. Ali podemos ver que o atendimento total prestado a residentes em Recife ficou ligeiramente abaixo do que seria a taxa média brasileira de atendimentos por habitante multiplicada pela população de Recife e ligeiramente acima do esperado pela taxa média de Pernambuco. A participação do prestador em questão se manteve como uma fração significativa deste atendimento, sem nunca ter atingido o atendimento total.

Dados mostrados nas folhas de fatos produzidas pelo INFOSAS podem ser verificados de forma independente por qualquer usuário, utilizando por exemplo o aplicativo Tabwin distribuído pelo DATASUS⁶.

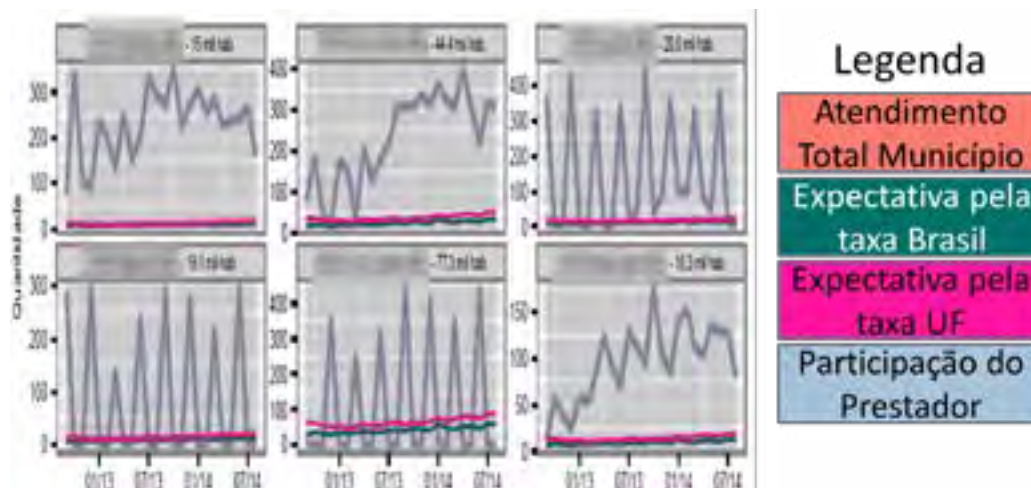


Figura 3: Um exemplo de anomalia estatística no atendimento por um mesmo prestador a residentes em municípios

Alguns fatos podem evidenciar a ocorrência de anomalias estatísticas, como é o caso mostrado na Figura 3: Um exemplo de anomalia estatística no atendimento por um mesmo prestador a residentes em municípios. Ali podemos observar que o número de atendimentos a todos os seis municípios mostrados está muito acima do esperado pelas taxas brasileiras e do estado, e a superposição das curvas do atendimento total do município e da participação do prestador indicam que estes atendimentos foram praticamente monopolizados pelo prestador em questão.

Outra anomalia possível é o valor excessivo cobrado em uma Autorização de Internação Hospitalar (AIH), como mostrado na Figura 4: Exemplo de anomalia estatística no valor de internações hospitalares de um prestador. O eixo horizontal mostra o tempo, marcado em meses, de junho de 2012 a maio de 2014. O eixo vertical mostra o valor médio pago por procedimento realizado no mês no alvo escolhido pelo usuário. Existem duas curvas. Uma, em verde, mostra o valor médio no Brasil. A outra, em azul, mostra o valor médio praticado pelo estabelecimento sob análise. Pequenos números em cada mês junto à curva azul mostram o número de procedimentos sobre os quais a média foi calculada. Claramente, o valor médio neste estabelecimento e neste alvo estão muito acima do valor típico no Brasil.

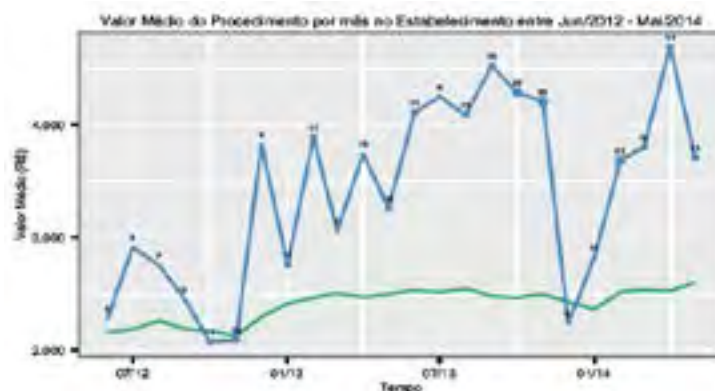


Figura 4: Exemplo de anomalia estatística no valor de internações hospitalares de um prestador

Anomalias encontradas pelo INFOSAS são apenas o que o nome indica: anomalias estatísticas, que podem sim ter sido provocadas por fraudes, mas que também podem resultar de processos corretos de ocorrência pouco frequente, como mutirões ou transferências de serviços entre prestadores, ou de informação incorreta nos registros de atendimentos, como o lançamento do endereço do hospital ao invés do endereço do paciente. Taxas de cobertura com valores muito acima do esperado pela taxa brasileira podem também simplesmente resultar de uma má distribuição do atendimento no país, com os municípios com atendimento anômalo sendo aqueles poucos com o atendimento correto. Cabe ao usuário do INFOSAS analisar cada caso para chegar à interpretação correta ou buscar mais informações diretamente com o prestador dos serviços ou gestor.

Com os devidos cuidados, anomalias estatísticas são informações de alto interesse para gestores. Folhas de fatos podem ser elementos de decisão para o lançamento por gestores de auditorias ou de outros procedimentos para uma apuração mais detalhada das causas das anomalias estatísticas.

Examinar todas as folhas de fatos procurando por indicações de anomalias é, entretanto, inviável. Temos aproximadamente 5000 alvos, 6000 prestadores e, se considerarmos somente 3 anos de produção, com janelas de 12 meses para os 36 meses, temos mais de 1 bilhão de folhas a serem examinadas. É aí que o INFOSAS mostra o seu valor, utilizando algoritmos que procuram capturar a ocorrência de anomalias, e produzindo *scores* numéricos que permitem ordenação e priorização das folhas de fatos.

O INFOSAS permite também que seu usuário restrinja sua análise a um determinado estado, a uma região de saúde, ou a um município, a um mês determinado e a uma extensão da janela de tempo onde os dados são examinados, e também a um alvo ou conjunto de alvos, conforme mostra a Figura 5: Filtros para seleção de alvo de mineração e de recorte geográfico para a localização de prestadores.

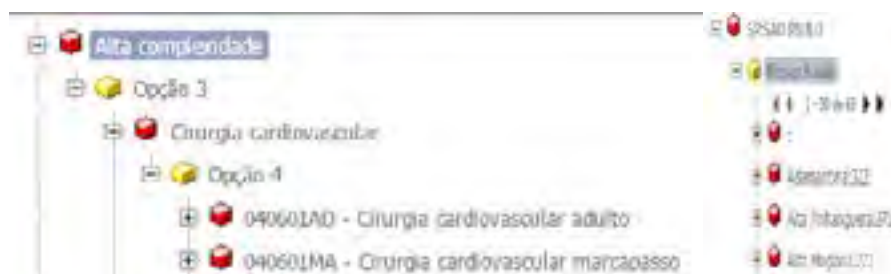


Figura 5: Filtros para seleção de alvo de mineração e de recorte geográfico para a localização de prestadores

Feita a seleção, o INFOSAS produz um relatório com as folhas de fatos com os 100 maiores *scores* de anomalia. A Figura 6: Relatório com as 100 folhas de fatos com maiores *scores* mostra uma parte de um relatório para o conjunto de alvos “030305 Tratamento de doenças do aparelho da visão”,

restrito à região Nordeste, com janela de tempo de análise de 24 meses terminando em agosto de 2014. Obscurecemos informações sobre municípios e estabelecimentos por não estarmos autorizados a divulgar esses dados pelo Ministério da Saúde.

Rank	Processamento	Descrição	Valor	Score	Valor	Score
104	ES - Procedimento clínico	T003 - Tratamento de doenças de aparelho de visão (O3000)	104	0,90	0,000	2.477,207.00
103	ES - Procedimento clínico	T003 - Tratamento de doenças de aparelho de visão (O3000)	103	0,88	0,000	223,707.27
102	ES - Procedimento clínico	T003 - Tratamento de doenças de aparelho de visão (O3000)	102	0,87	0,000	2.044,200.00
101	ES - Procedimento clínico	T003 - Tratamento de doenças de aparelho de visão (O3000)	101	0,85	0,000	186,000.00
100	ES - Procedimento clínico	T003 - Tratamento de doenças de aparelho de visão (O3000)	100	0,84	0,000	170,000.00

Figura 6: Relatório com as 100 folhas de fatos com maiores escores

No restante deste artigo, vamos descrever os princípios e diretrizes adotados pelo INFOSAS e os principais algoritmos usados para a detecção das anomalias. Como os dados são massivos, o armazenamento, compressão e gestão de dados devem ser cuidadosamente concebidos. Igualmente importantes, o anonimato e segurança dos dados devem ser garantidos. Estes aspectos do INFOSAS não são tratados neste artigo. Nós vamos focar nos principais algoritmos desenvolvidos para detectar anomalias e na sua combinação para gerar um escore unificado para cada provedor de serviços de saúde. Em seguida, apresentamos alguns resultados obtidos com o INFOSAS e os ganhos que ele pode trazer para o SUS. Concluimos com algumas perspectivas sobre o desenvolvimento futuro do INFOSAS.

Métodos

Princípios e Diretrizes

O INFOSAS foi construído com base em certos princípios e diretrizes definidas a priori ou que fomos aprendendo e assimilando à medida que trabalhamos no projeto. Desde o início sabíamos que não haveriam exemplos rotulados de anomalias ou fraudes para nos guiar na construção de regras de classificação de novos casos. O conhecimento de não conformidades identificadas por auditoria do Departamento de Regulação, Avaliação e Controle de Sistemas (DRAC) do Ministério da Saúde permitiria a confirmação de anomalias eventualmente detectadas pelo INFOSAS. Infelizmente, esta informação não está disponível e isto nos jogou no mundo da análise de classificação não-supervisionada, uma seara mais difícil e mais imprecisa que o mundo da classificação supervisionada^{7,8}. Decidimos focar as técnicas de análise nas séries temporais de procedimentos realizados nos municípios de residência dos pacientes por cada estabelecimento. Um segundo elemento que nos surpreendeu no início foi a heterogeneidade de padrões estatísticos nestas séries temporais. Nenhum modelo conhecido de séries temporais, tais como modelos ARIMA ou modelos de espaços de estados⁹, seria capaz de dar conta de toda a variação que observávamos nas séries.

Percebemos que a conclusão alcançada em outros domínios também é válida no nosso caso. A saber, que nenhum algoritmo de detecção poderia detectar sozinho todos os tipos de fraude. Por isto, o INFOSAS usa muitos algoritmos simples de detecção em vez de um único algoritmo complexo. Cada

um desses algoritmos simples gera um escore de discrepância. Quanto maior seu valor, mais anômalo é o estabelecimento no aspecto capturado pelo indicador.

Usar vários algoritmos simples implica no problema de combiná-los apropriadamente. Vários algoritmos de detecção podem gerar um alerta sobre um mesmo estabelecimento de saúde, e esses alertas são combinados em um único escore final associado ao estabelecimento. Esta combinação deveria levar em conta a importância diferenciada dos algoritmos.

Outro aspecto óbvio é a necessidade de usar métodos automáticos por causa do volume massivo de dados. Menos evidente é a necessidade de usarmos métodos robustos de análise. Na análise de dados, métodos robustos são aqueles que não são sensíveis a pontos aberrantes (ou *outliers*) ou a erros grosseiros. Por exemplo, a mediana de um conjunto de dados é sempre preferível ao cálculo da média aritmética, por este critério. As duas medidas fornecem uma estimativa de valor central, mas a mediana não é afetada pela presença de até 50% de dados muito diferentes dos demais, enquanto a média aritmética pode ser drasticamente modificada.

Uma diretriz fundamental que norteia o INFOSAS desde o início é o custo muito elevado de falsos positivos. Após identificar de forma automática um estabelecimento, ele deve sofrer um escrutínio detalhado por parte de analistas. O INFOSAS detecta discrepâncias estatísticas que podem ser efeitos colaterais de irregularidades, mas que também podem resultar de processos corretos de ocorrência pouco frequente. Exemplos possíveis incluem os mutirões ou transferências de serviços entre prestadores de serviços que provocam eclosões de procedimentos similares num curto espaço de tempo num mesmo local. Comparando com a história deste local, esta eclosão é uma anomalia e acaba sendo captada pelo INFOSAS. Outro problema é o lançamento de informação incorreta, como o lançamento do endereço do hospital ao invés do endereço do paciente. Cabe ao gestor examinar evidências e eventualmente decidir por alguma ação. De qualquer forma, esta é uma atividade que consome muito tempo e recursos humanos muito qualificados. Caso o analista decida que as evidências são grandes, uma auditoria presencial aumenta os custos do processo deflagrado pelo INFOSAS. Dessa forma, o efeito de falsos positivos pode ser muito grande, podendo até mesmo inviabilizar o sistema se eles forem muito frequentes.

Algoritmos

Procuramos detectar dois tipos de discrepâncias: um valor médio excessivo cobrado por procedimentos dentro de um alvo e um número excessivo na produção de um procedimento por parte de um estabelecimento. Analisamos as séries temporais de valor médio mensal por procedimento e de produção mensal em cada alvo desejado. Tais séries são calculadas por estabelecimento e por município de residência do paciente. Vários indícios de anomalias são analisados e, para cada um deles, um escore de discrepância é calculado. Estes escores possuem faixas de variação muito diferentes e precisam ser combinados.

Por falta de espaço para descrever em detalhes todos os escores, vamos apresentar a seguir a definição de dois deles, junto com exemplos ilustrativos de seu uso. A seguir, mostramos como combinamos os escores.

Desvio da mediana de valor/produção

O objetivo deste algoritmo é comparar a produção (ou o valor médio mensal cobrado) pelo estabelecimento com sua própria série histórica de produção (ou de valor médio). Se houver uma mudança muito abrupta na série, o algoritmo atribui um escore de discrepância ao estabelecimento. Mais especificamente, em cada mês i e para cada estabelecimento l , nós calculamos

$$escore_{li} = \frac{t_{li} + 1}{mediana_{li} + 1}$$

onde t_{li} é a quantidade de interesse (produção/valor médio) para o l -ésimo estabelecimento no i -ésimo mês, $mediana_{li}$ é a mediana dos últimos m meses para a quantidade de interesse do l -ésimo estabelecimento no i -ésimo mês.

A Figura 7 mostra o uso deste algoritmo com dados de certo alvo. A linha azul mostra a série mensal do estabelecimento. A linha vermelha é a mediana dos últimos 12 meses desse mesmo estabelecimento. Assim, a linha vermelha captura uma tendência histórica do estabelecimento, mas não é afetada por anomalias presentes nos últimos seis meses a partir do mês de referência. A produção deste estabelecimento no mês de março de 2012 foi de 34 procedimentos. A mediana dos últimos 12 meses (a partir de março de 2012) é igual a 11. Desta forma, neste mês e neste estabelecimento, o escore associado com este algoritmo é igual a $(34+1)/(11+1) = 2,92$.

Produção do município

O objetivo deste algoritmo é estabelecer um limiar para a produção do estabelecimento exercida nos municípios em que ele atua tomando como base a produção brasileira por 100 mil habitantes. Analisamos a série temporal de produção realizada por todos os estabelecimentos nos residentes de um município. Caso este município tenha uma produção acima do limiar, o algoritmo atribui um escore de discrepância para ele. Em seguida, o escore é repartido proporcionalmente aos estabelecimentos que produziram nesse município.

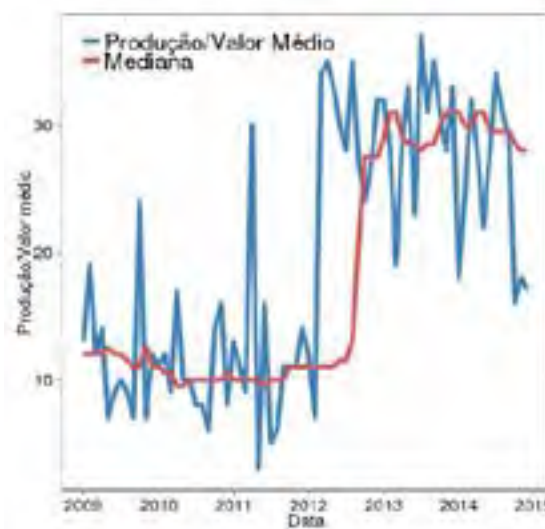


Figura 7: Ilustração do algoritmo do desvio da mediana

O escore do estabelecimento é obtido somando-se os escores que ele obteve em cada município em que atuou. Mais formalmente, o escore de produção do município é calculado como uma soma acumulada nos últimos 12 meses:

$$escore_{li} = \sum_{j=i-11}^i dif_{lj}$$

onde $dif_{li} = \max\{0, tBayes_{li} - limiar_i\}$ e $limiar_i = k_3 * taxaBrasil_i$. O valor de $taxaBrasil_i$ é a taxa de produção mensal brasileira por 100 mil habitantes no mês i , $tBayes_{li}$ é a taxa bayesiana empírica por 100 mil habitantes no mês i para o município l e k_3 é uma constante previamente definida. Em nossos estudos, temos usado $k_3 = 3$. A taxa bayesiana empírica é uma técnica estatística para calcular taxas e razões que não é afetada pela flutuação de pequenas populações¹⁰. É bastante conhecido

o problema de calcular taxas de eventos raros em populações pequenas, quando um ou dois eventos a mais podem fazer a taxa flutuar brutalmente. A Figura 8 ilustra o uso deste algoritmo considerando 6 municípios em que um certo estabelecimento atua.

A linha azul representa toda a produção realizada nos residentes de cada município enquanto a linha vermelha denota um limiar calculado a partir da taxa brasileira, da população do município, acrescida de uma enorme margem de segurança (três vezes o que se espera em média no município). Assim, comportamentos como aqueles verificados nos municípios 1 e 2 são muito atípicos, desviando-se muito do que é esperado.

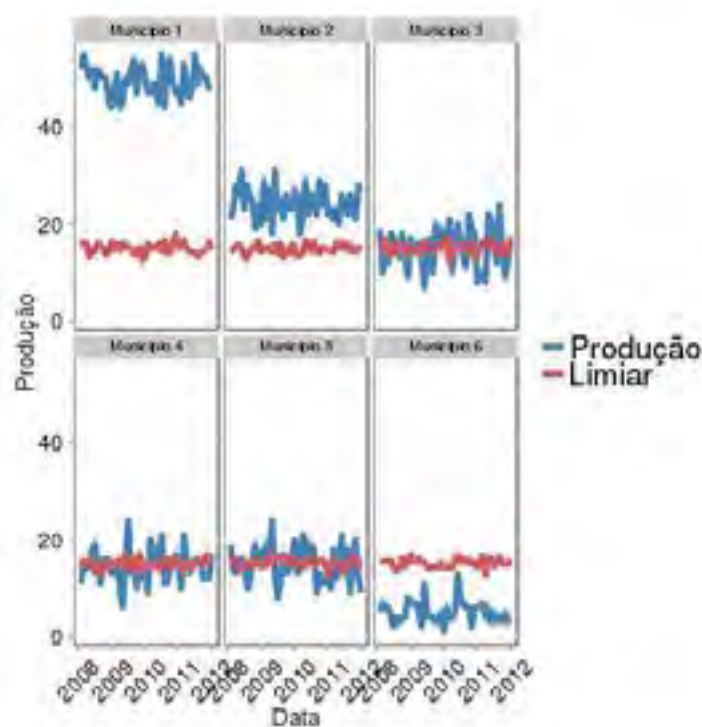


Figura 8: Ilustração do algoritmo da produção do município.

O score pode ser igual a zero caso a linha azul esteja abaixo do limiar nos últimos 12 meses. Cada estabelecimento prestando serviços às populações desses municípios recebe uma parcela do score de anomalia calculado em cada município. Esta parcela é proporcional ao volume de serviços prestado pelo estabelecimento aos residentes daquele município.

Combinando os escores

INFOSAS possui uma série de algoritmos como estes que apresentamos anteriormente, cada um deles procurando captar algum tipo de discrepância estatística na série temporal da produção ou do valor médio associado com um estabelecimento. Os escores devem ser combinados para criar um score unificado. Como as escalas numéricas dos diferentes escores são muito distintas, esta unificação não é óbvia.

Nosso procedimento básico foi transformar cada um dos escores de forma que eles estejam no intervalo entre 0 e 1. Quanto mais próximo do valor 1 mais anômalo é o estabelecimento naquele score. Para obter este valor entre 0 e 1, ajustamos de forma robusta uma distribuição de probabilidade gaussiana $N(\mu, \sigma^2)$ aos dados estatísticos de cada score. Após colocarmos os escores numa mesma escala, podemos combiná-los para criar um único número. Para isto, atribuímos pesos para cada algoritmo de acordo a capacidade esperada de detectar anomalias. Este peso é subjetivo e está associado com a importância que atribuímos a cada um dos algoritmos. O score unificado é dado por

$$escore_{jk}^* = escore_{ji} * w_j \quad \sum_{j=1}^k w_j = 1$$

onde w_j é o peso do j -ésimo *escore* e $escore_{ji}$ é o j -ésimo *escore* padronizado do i -ésimo estabelecimento no i -ésimo mês. Estamos estudando formas de validar a escolha desses pesos.

Resultados e discussão

um dos resultados mais importantes do INFOSAS é a sua capacidade de análise do volume total de produção de serviços de saúde em busca de discrepâncias considerando cada prestador. Esta capacidade estará ao alcance das mãos dos analistas do Ministério da Saúde, dos estados e dos municípios, numa interface simples e amigável. Além desse uso em microescala, podemos também apresentar alguns resultados que o INFOSAS poderá trazer quando adotado. Mostramos nesta seção alguns números indicando que um esforço pequeno, concentrado e guiado pelo INFOSAS pode ser recompensado amplamente. Em particular, obtemos uma estimativa do valor em excesso que o SUS está dispendendo.

Analizamos usando os dados do ano de 2013. Foram movimentados um total de 19.9 bilhões de reais para realizar 789 milhões de atendimentos médicos. Deste total, numa estimativa muito conservadora, consideramos que 414 milhões foram gastos em excesso, ou 2% do total, divididos entre 350 milhões em alvos SIH e 64 milhões em alvos SIA. Analizamos 269 alvos disjuntos em 5570 municípios e mais de 23 mil prestadores de saúde, num total de mais de 1.5 milhões de séries temporais.

Encontramos 6811 prestadores com valores considerados excedentes e discrepantes. Este grupo de prestadores é bastante desigual. O valor excedente concentrado nos 100 prestadores mais críticos é de 210 milhões de reais, ou 51% do total excedente estimado. Na verdade, a concentração é bastante elevada: os top 5 prestadores são responsáveis por 9% do valor excedente, os top 10 por 15% e os top 20 por 24%. Apenas estes 20 prestadores mais críticos significam um valor excedente de 98 milhões de reais.

Para chegar a estes números, consideramos a série temporal mensal da taxa de cobertura de cada um dos 5570 municípios para cada um dos 269 alvos. A taxa de cobertura é a simplesmente taxa per capita de realização de atendimentos no mês. Para evitar flutuações estatísticas não associadas com desvios reais, aplicamos a correção bayesiana empírica de pequenas populações a estas taxas. Verificamos que distribuição log-normal fornece um excelente ajuste a estas taxas. Considerando apenas os municípios *outliers* em cada alvo, repartimos o valor excedente acima de um limiar extremamente elevado entre os prestadores que atenderam ao município-alvo em questão de forma *pro-rata*.

Conclusão

O sistema INFOSAS já foi apresentado em diversos foros nacionais (veja por exemplo um vídeo da Oficina de Controle e Avaliação, realizado em Brasília, em agosto de 2015, disponível no endereço https://www.youtube.com/watch?v=vIaR_Q7T-Us), e em todos eles, seus resultados foram julgados muito interessantes por especialistas em saúde pública. Por isso, acreditamos que o sistema INFOSAS pode ser utilizado no processo de indicação de casos para auditoria, melhorando a sua qualidade e reduzindo a frequência de auditorias desnecessárias.

Os próximos passos passam certamente pela realização de verificações *in loco*. Isto nos permitiria, por um lado, refinar os algoritmos já utilizados. Por outro lado, ao gerar uma amostra de casos corretamente rotulados como *anômalos* e *não-anômalos*, possibilita a utilização de métodos de aprendizado supervisionado. Estas técnicas de classificação supervisionadas são bastante poderosas e permitem uma melhor caracterização da tipologia de fraudes. No momento atual, a pedido do MS, estamos desenvolvendo um curso à distância de utilização do INFOSAS. Isto vai permitir o seu uso diretamente por gestores de todas as instâncias do SUS.

Agradecimentos

Este trabalho recebeu o apoio financeiro do Ministério da Saúde, FAPEMIG e CNPq. Os autores agradecem a todos os membros do projeto INFOSAS. Em particular, agradecemos a Maria Helena Brandão, Raquel Ferreira, Douglas Mesquita, Milton Pífano, Larissa Castro dos Santos, Tomas Schweizer, Fabiana Pereira Peixoto e Leticia Santos Neto. Agradecemos também ao então Secretário de Atenção Primária do Ministério da Saúde, Dr. Helvécio Magalhães, que em 2013 foi quem primeiro vislumbrou a possibilidade de um sistema automático de detecção de anomalias para o SUS.

Referências

- [1] Home - The NHCAA [Internet]. [citado 30 de maio de 2016]. Recuperado de: <https://www.nhcaa.org/>
- [2] SIA - DATASUS [Internet]. [citado 30 de maio de 2016]. Recuperado de: <http://datasus.saude.gov.br/sistemas-e-aplicativos/ambulatoriais/sia>
- [3] SIHSUS - DATASUS [Internet]. [citado 30 de maio de 2016]. Recuperado de: <http://datasus.saude.gov.br/sistemas-e-aplicativos/hospitalares/sihsus>
- [4] Cadastro Nacional de Estabelecimentos de Saúde [Internet]. [citado 30 de maio de 2016]. Recuperado de: <http://cnes.datasus.gov.br/>
- [5] SIGTAP - Sistema de Gerenciamento da Tabela de Procedimentos, Medicamentos e OPM do SUS [Internet]. [citado 1 de fevereiro de 2014]. Recuperado de: <http://sigtap.datasus.gov.br/tabela-unificada/app/sec/inicio.jsp>
- [6] DATASUS [Internet]. [citado 30 de maio de 2016]. Recuperado de: <http://www2.datasus.gov.br/DATASUS/index.php?area=060805>
- [7] Zaki M, Meira Jr.W. Data Mining: Fundamental Concepts and Algorithms. New York: Cambridge University Press; 2014.
- [8] Murphy K. Machine Learning: a probabilistic perspective. Cambridge: The MIT Press; 2012.
- [9] Shumway RH, Stoffer DS. Time Series Analysis and Its Applications: With R Examples. 3rd. ed. New York: Springer; 2011.
- [10] Assunção RM, Barreto SM, Guerra, HL, Sakurai E. Mapas de taxas epidemiológicas: uma abordagem Bayesiana. Cadernos de Saúde Pública, 1998, 14(4); 713-23.

Contato

Autor correspondente
Renato M. Assunção
Professor Titular, Departamento de Ciência da
Computação, Universidade Federal de Minas
Gerais
assuncao@dcc.ufmg.br