

# Programação genética aplicada no processo de previsão: um estudo de caso aplicado em chamadas de uma central de teleatendimento

Cláudio Lúcio do Val Lopes, Gustavo Henrique Passini Santos e  
Flávio Vínicius Cruzeiro Martins

CEFET-MG Av. Amazonas, 7675 - Nova Gameleira, Belo Horizonte - MG,  
30510-000, Brasil

{claudiolucio,passini.gustavo}@gmail.com  
flaviocruzeiro@decom.cefetmg.br

**Resumo** Este trabalho apresenta um estudo de caso utilizando a programação genética como uma proposta para obter um modelo de previsão de chamadas para centrais de teleatendimento. É apresentado também um modelo clássico utilizado em séries temporais com um ajuste para as centrais de teleatendimento. Um processo de extração de características da série é apresentado e utilizado na programação genética. Os resultados do estudo de caso indicam que a programação genética é viável com resultados superiores ao modelo clássico. Melhorias e trabalhos futuros são indicados no intuito de explorar esta abordagem de computação evolucionária para séries temporais em previsão de chamadas.

## 1 Introdução

A forma de interação entre empresas e seus clientes tem cada vez mais tomado o formato de centrais de atendimento com várias modalidades de interação, dentre elas o teleatendimento. Uma central de teleatendimento é um conjunto de recursos (tipicamente computadores, equipamentos de telecomunicações e pessoal para atendimento) que está habilitado a prover serviços via telefone [4].

A necessidade de estabelecer e manipular um grande número de contatos com os clientes tem feito com que empresas busquem melhorar seus processos e serviços. Muitas empresas possuem departamentos para lidar com esta temática: “*call centers*”, “*contact centers*” ou centro de interações com clientes [16].

As centrais de teleatendimento buscam atingir uma determinada qualidade (utilizando os níveis de serviço para medição) sujeitos a recursos finitos (orçamentos, atendentes, linhas telefônicas, equipamentos). Neste contexto existem vários tipos de questões relacionadas a tomada de decisão [4]: estratégico, tático, planejamento e em tempo real.

Uma das questões relacionadas ao tipo de decisão de planejamento é a previsão do número de chamadas. A quantidade de chamadas nas centrais de teleatendimento são observadas em intervalos diários e intradiários (de hora em hora ou até mesmo de 15 em 15 minutos). Este problema também é conhecido como

previsão de demanda, e é um tema importante para a operação das centrais: programação de trabalho de atendentes com escalas e turnos, nível de serviço, contratações de mão de obra, realização de treinamentos dentre outros.

A programação genética é uma técnica automática de programação que propicia a evolução de programas. Alguns trabalhos utilizam o paradigma da programação genética em conjunto com a regressão simbólica para lidar com o tema de previsão de séries temporais [3], [10] e [1]. Este tipo de abordagem é interessante para a automatização do processo de geração de boas soluções e tem mostrado resultados em algumas áreas de aplicação [6] e [15].

O objetivo deste trabalho é aplicar o paradigma da programação genética para um estudo de caso específico de previsão de quantidade de chamadas intradiárias em uma central de teleatendimento e comparar tal abordagem com resultados de um modelo clássico. O restante deste trabalho está organizado da seguinte forma: na próxima seção será apresentada a ideia geral da programação genética e regressão simbólica. O problema da previsão de chamadas será explicitado na seção 3, assim como um breve resumo dos modelos utilizados para previsão de chamadas: modelo clássico e modelo da programação genética. A seção seguinte explica os experimentos realizados, a partir de um estudo de caso (utilizando uma série real), detalhando os modelos utilizados e resultados obtidos com a comparação. Finalmente, na última seção, serão feitos alguns comentários finais e observações.

## 2 Programação genética

A programação genética (PG) é uma técnica de computação evolucionária que automaticamente resolve problemas, sem a necessidade que o usuário antecipadamente especifique a forma ou estrutura da solução. É um método sistemático e independente de domínio para que computadores possam resolver problemas a partir de uma definição de alto nível do que precisa ser feito [14].

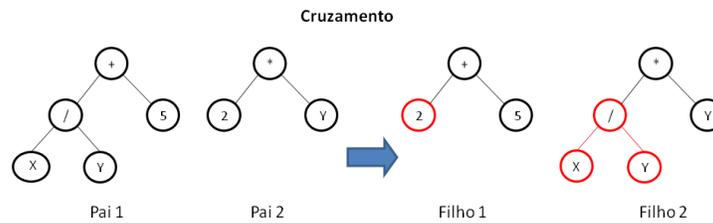
A programação genética busca evoluir programas de computadores. A abordagem é populacional e evolucionista permitindo que geração após geração, de forma estocástica, haja a transformação dos programas em “versões” que buscam ser melhores de acordo com o critério adotado. Programas criados por meio da PG são geralmente caracterizados por árvores sintáticas, em que os nós simbolizam a representação do problema e a árvore o fluxo de execução do programa.

O processo de desenvolvimento de uma PG inicia-se com a definição dos nós que formarão as árvores sintáticas, dividindo-os em dois grupos: Nós não Terminais (funções com argumentos, por exemplos) e nós terminais (tipicamente variáveis, constantes e funções sem argumentos). Como parâmetros da PG podem ser informados, por exemplo: a profundidade máxima da árvore, número máximo de nós e a quantidade de programas em cada geração [14].

O processo de inicialização da população é realizado com a geração aleatória de programas. Há alguns métodos de criação de indivíduos: a técnica *full* gera árvores que sempre atingem sua profundidade máxima; a técnica *grow* permite árvores com várias alturas e configuração de nós, respeitando os parâmetros; e

a técnica *Ramped half-and-half* que gera metade das árvores com a técnica *grow* e o restante com a técnica *full* [15].

Geração após geração os programas são avaliados recebendo um valor de aptidão (a avaliação é realizada seguindo as regras do problema em questão). Processos de seleção são aplicadas em cada geração. Como exemplo, pode-se citar: I) seletor aleatório, em que a seleção ocorre de forma estocástica; II) torneio, no qual uma pré-seleção aleatória de  $x$  programas é realizada, sendo selecionado o que obtiver melhor valor de aptidão dentre os participantes do torneio; III) seletor probabilístico, em que uma probabilidade de escolha é atribuída a cada programa de acordo com sua aptidão, seguido de um sorteio que leva em conta tal probabilidade atribuída [14].



**Figura 1.** Exemplo de um cruzamento de ponto único entre árvores sintáticas genitoras.

O cruzamento entre programas também é utilizado no processo evolutivo, esperando que as melhores características de cada programa sejam compartilhados na sua prole. Como exemplo de cruzamento considera-se o cruzamento de ponto único (um nó de cada programa é selecionado realizando a troca entre os mesmos para criação de um novo programa) e o cruzamento de subárvore (um nó de cada programa é selecionado e as subárvores contendo esses nós são trocados para gerar um novo indivíduo - veja um exemplo na Figura 1)



**Figura 2.** Exemplo de uma mutação em uma árvore sintática.

A mutação também é aplicada ao processo evolutivo, buscando manter a diversidade da população. Como exemplo, cita-se a mutação de um nó e de uma

subárvore, em que essas subestruturas são completamente alteradas, formando um novo programa (veja um exemplo na Figura 2).

Finalmente, os critérios de parada são verificados, caso sejam atingidos, a solução (programa) é entregue como saída. Caso contrário, a substituição geracional é realizada.

### 3 Previsão de chamadas

As previsões de chamadas em centrais de teleatendimento devem levar em consideração três aspectos [2] : I) previsões do total de chamadas ou previsões por tipo de chamadas (chamadas para cancelamento de serviço, para novos clientes, para atendimentos especializados); II) dados disponíveis para o modelo de previsão, por exemplo, em algumas centrais de teleatendimento os níveis de serviço são monitorados de 15 em 15 minutos, de forma horária ou diária; e III) o intervalo de previsão gerado pelo modelo, necessário para tomada de decisão. Este item está relacionado ao tipo de questão que o modelo de previsão deve responder. Para questões táticas será necessário, por exemplo, uma previsão para as próximas duas semanas com previsões horárias, para questões de planejamento, talvez seja necessário prever o próximo mês com previsões diárias, por exemplo.

Modelos de previsão diários e intradiários com intervalos de previsão horários são apresentados em [17]. Uma abordagem para reduzir a dimensionalidade dos dados intradiários é aplicada. Além da redução da dimensionalidade, uma transformação nos dados é aplicada, e modelos Gaussianos mistos de previsão são utilizados. Um estudo comparativo utilizando dados reais é apresentado em [8]. A mesma transformação da variável, quantidade de chamadas, utilizada em [17] é também aplicada neste artigo. A previsão por tipo de chamadas é outro aspecto tratado com alguns modelos mistos apresentados.

Na presença de dados intradiários alguns modelos são utilizados na literatura: ARMA, alisamento exponencial e modelos Bayesianos. Uma revisão de trabalhos que lidam com modelos de previsão para centrais de teleatendimento (com comparações) é apresentada em [9].

#### 3.1 Modelo ARMA - Auto Regressivo com médias móveis

A quantidade de chamadas pode ser representada por uma variável aleatória com distribuição de Poisson [11]. A proposta do modelo clássico apresentada neste trabalho utiliza regressão dinâmica aplicados à média da série. Os modelos de regressão dinâmica consideram a distribuição da variável aleatória como Gaussiana. Uma transformação na quantidade de chamadas é aplicada para lidar com natureza da distribuição de Poisson [18].

A quantidade de chamadas é uma variável aleatória discreta. Para lidar com a natureza discreta desta variável e para estabilizar a variância, uma alternativa é o método apresentado em [18]. Considere  $N_{dk}$  como a quantidade de chamadas e  $\lambda_{dk}$  a taxa de chegada das chamadas no dia  $d$ ,  $d \in \{1, \dots, D\}$ , e da hora  $k$ ,  $k \in \{1, \dots, K\}$ . Observa-se que  $N_{dk} \approx Poisson(\lambda_{dk})$  [18] e assume-se que a seguinte

transformação  $y_{dk} = \sqrt{N_{dk} + 1/4}$  possui uma média de  $\sqrt{\lambda_{dk}}$  e variância de  $1/4$ . Em [17] é apresentada a propriedade assintótica, ou seja, quando  $\lambda \rightarrow \infty$  então  $y_{dk}$  é aproximadamente normal. Desta forma os modelos Gaussianos utilizam esta transformação, necessária, para que não haja nenhum tipo de viés nos coeficientes estimados.

As previsões tratadas pelos modelos Gaussianos irão tratar a variável  $y_{dk}$ . O estimador do número de chamadas será tratado como  $\hat{N}_{dk} = \hat{y}_{dk}^2 - 1/4$ . Esta proposta é utilizada em: [8] e [17].

Na modelagem de auto-correlação temporal os modelos Gaussianos do tipo auto-regressivo e de médias móveis ARMA são adequados [13]. Os modelos ARMA adotam a nomenclatura  $p$  para indicar a ordem do componente auto-regressivo e  $q$  para indicar a ordem do componente de médias móveis. Os modelos propostos neste artigo também possuem efeitos fixos, derivados dos dias e horas. Um modelo  $ARMA(p,q)$  é definido pela equação 1.

$$X_t - \mu = \phi_1(X_{t-1} - \mu) + \dots + \phi_p(X_{t-p} - \mu) + \epsilon_t - \theta_1\epsilon_{t-1} - \dots - \theta_q\epsilon_{t-q} \quad (1)$$

Em que  $\mu, \phi_1.. \phi_p, \theta_1.. \theta_q$  são parâmetros reais e  $\epsilon_t \approx Normal(0, \sigma^2)$  e com as seguintes características: temporalmente homogêneo, estacionário e sem dependência temporal. A média do processo é representado por  $\mu$ . Para este modelo devem ser observadas as condições de estacionariedade dos componentes auto-regressivos e a condição de invertibilidade dos componentes de médias móveis.

### 3.2 Modelo de programação genética

Alguns trabalhos aplicam computação evolucionária para executar previsão de series temporais [10],[1] e [7]. Estes trabalhos buscam criar um modelo de previsão utilizando a temática de populações através da PG. Os programas são criados utilizando a ideia de regressão simbólica em que a estrutura das funções é flexível e pode ser linear ou não linear ([3] e [12]).

Em outros tipos de regressão a estrutura da função é pré determinada. Na regressão simbólica a estrutura da função é flexível e obtida por algum processo de alteração de sua estrutura, seus termos e coeficientes. A utilização de computação evolucionária propõe-se a modelar características dos dados, principalmente características não lineares. Várias tipos de séries temporais apresentam características não lineares em seu comportamento.

Em [10] são utilizados modelos clássicos para modelar o componente linear das séries e um algoritmo genético para tratar o componente não linear. Já em [7] um algoritmo genético é aplicado em series temporais para detectar padrões nas séries de dados sobre terremotos. Em [1] a programação genética é utilizada para classificar séries temporais. Já em [6] um conjunto de experimentos é realizado para comparação da PG com métodos como ARIMA e alisamento exponencial. Os autores ainda propõem um método que seleciona as previsões a partir de vários algoritmos de PG para um mesmo problema, apresentando resultados superiores quanto aos modelos comparados.

Para o caso de previsão de séries de chamadas em centrais de teleatendimento não encontrou-se referência que fez uso do PG para realizar tais previsões.

## 4 Estudo de caso e resultados

Em [11] é apresentado um estudo sobre uma base de dados pertencente a uma instituição financeira do Estado de Israel. A base de dados pode ser obtida em: <http://ie.technion.ac.il/serveng/callcenterdata/index.html>. Este trabalho obteve permissão do Prof. Avishai Mandelbaum para utilização da base de dados. A base de dados possui 444.448 chamadas recebidas. A central de teleatendimento oferece vários tipos de serviços e seu horário de funcionamento é de 07:00 às 24:00. A quantidade de chamadas é ,em média, de 30.000 a 40.000 chamadas por mês (chamadas de clientes para falar com atendentes).

A base de dados utilizada para fazer o ajuste dos modelos de previsão apresenta 3.612 observações oriundas de um total de 356.638 chamadas da base original. Cada observação contém a data (dia/mês/ano) e hora do dia de 09:00 às 22:00.

O objetivo da previsão do estudo de caso é para lidar com questões de tomada de decisão táticas relacionadas ao planejamento das centrais de teleatendimento. Para tal é necessário que se trabalhe com um período de previsões, que para este estudo de caso, foi de 2 semanas (compreendido entre 12 até 23 de dezembro).

### 4.1 Modelo ARMA proposto

O modelo ARMA proposto utiliza fatores fixos das horas e dias em conjunto com um modelo  $ARMA(2,2)$ . Os valores  $p$  e  $q$  do modelo  $ARMA(2,2)$  foram obtidos através da análise das funções de autocorrelação e de autocorrelação parcial. A equação deste modelo é apresentada em (2), ajuste da série transformada. Observe que por questões de notação utilizou-se:  $N_{dk}$  como a quantidade de chamadas no dia  $d \in \{1, \dots, D\}$ , e da hora  $k, k \in \{1, \dots, K\}$  e  $y_{dk} = yt, t \in \{1, \dots, T\}$  e  $T = 3612$ . Os coeficientes  $\alpha_d$  e  $\beta_k$  são relativos às variáveis indicadoras para os dias dos meses e horas do dia, respectivamente. Também foram utilizadas algumas variáveis para tratar casos atípicos.

$$\hat{y}_t = \sum_{d=1}^D \alpha_d I_d + \sum_{k=1}^K \beta_k I_k + \Theta_0 + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} \quad (2)$$

### 4.2 Modelo de programação genética proposto

O modelo de programação genético adotado neste trabalho utiliza a PG para modelar os componentes lineares e não lineares da série do estudo de caso. Para a definição dos nós terminais foi utilizado um conjunto de constantes gerados na criação de cada candidato a solução. Este valores foram obtidos através de um processo de extração de características da própria série, como por exemplo:

desvio padrão, média, mediana e quartis da série, média e desvio padrão de chamadas por hora (com um total de 10 argumentos). Outras constantes aleatórias também foram inseridas com valores iniciais gerados a partir da distribuição de *Poisson* (o parâmetro  $\lambda$  utilizando foi a média geral da série para o período de ajuste/treinamento). Foram 32 constantes aleatórias disponíveis ao processo de evolução do programa

Além das constantes foi inserido também um grupo de argumentos oriundos da base de dados, representados pela defasagem da série e valores lógicos que indicavam a hora do dia em questão (total de 25 argumentos com esta finalidade).

Os nós não terminais foram definidos por meio de um grupo de funções de ponto flutuante como: soma, subtração, multiplicação, cosseno e seno. Um grupo de funções lógicas como: ‘se’, ‘e’, ‘ou’, igualdade e negação também foram utilizados na representação dos nós terminais.

Para a função *fitness* utilizou-se a raiz da média do erro quadrático (RMEQ) para avaliar a qualidade do ajuste dos modelos. A equação (3) é o RMEQ, na qual  $n$  é o número de amostras,  $\hat{y}_t$  é o valor fornecido pelo modelo para a  $i$ -ésima amostra e  $\bar{y}_t$  é a média dos valores de todas as amostras.

$$RMEQ = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_t - \bar{y}_t)^2} \quad (3)$$

O método de seleção adotado foi o torneio com três elementos (o elemento de melhor aptidão é selecionado). O método de cruzamento utilizado foi o ponto único. É importante ressaltar que a taxa de cruzamento foi de 0.8. Já o método de mutação utilizado foi o de um nó ou subárvore. A taxa de mutação foi de 0.1.

**Tabela 1.** Configurações com os parâmetros para o algoritmo da programação genética

Característica do algoritmo	Descrição
<i>Fitness</i>	Raiz da média do erro quadrático - RMEQ
Nós terminais	Constantes e argumentos
Nós não terminais (funções)	Algumas funções de ponto flutuante e funções lógicas
Tamanho máximo da Arvoré	De 3 até 30 níveis
Seleção	Torneio com 3 elementos
Cruzamento	Ponto único com taxa de 0.8
Mutação	Uniforme com taxa de 0.1
Condição de parada	400 gerações
Tamanho da população	100 indivíduos
Elitismo	Não, mas o DEAP mantém o “Hall of fame”

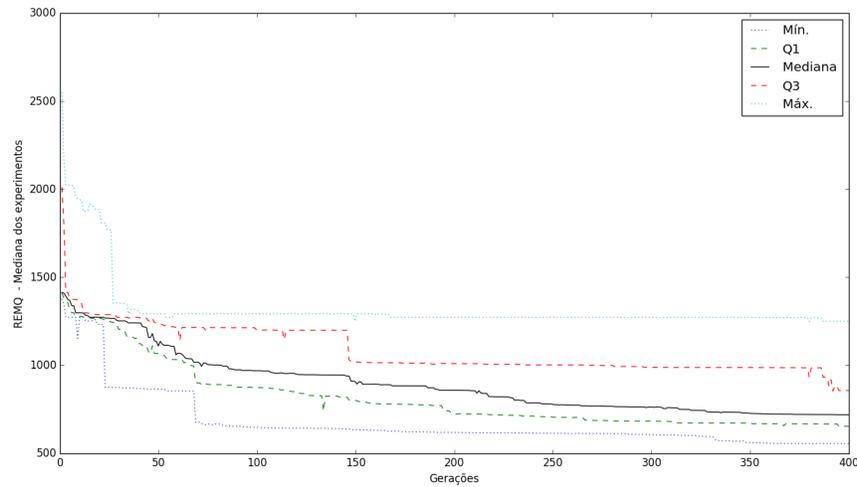
A solução aplicada mantém um “*Hall of fame*”, ou seja, os melhores indivíduos são armazenados e mantidos como soluções finais (não foi utilizado elitismo no processo de evolução). O critério de parada, o tamanho de árvore, e o número de níveis da árvore foram definidos observando o elevado tempo de

execução do processo evolucionário. Um resumo do algoritmo da PG criado para o estudo de caso pode ser visto na Tabela 1.

O modelo foi implementado utilizando um *framework* para algoritmos evolucionários na linguagem python chamado *Distributed Evolutionary Algorithm Python - DEAP* [5]. Não foi utilizada características para execução paralela e distribuída.

### 4.3 Resultados

Para o estudo de caso apresentado foram realizados dez experimentos para a PG. O tempo de execução do processo evolucionário foi de, aproximadamente, 10 horas (para um computador Intel com processador i5 e 16Gb de memória RAM). É possível observar pela curva de convergência apresentada na Figura (3) que a tática do elitismo não é utilizada (não observa-se uma descida linear no gráfico). Porém a mediana do erro médio quadrático vai diminuindo com o passar das gerações, satisfazendo o objetivo do estudo de caso.

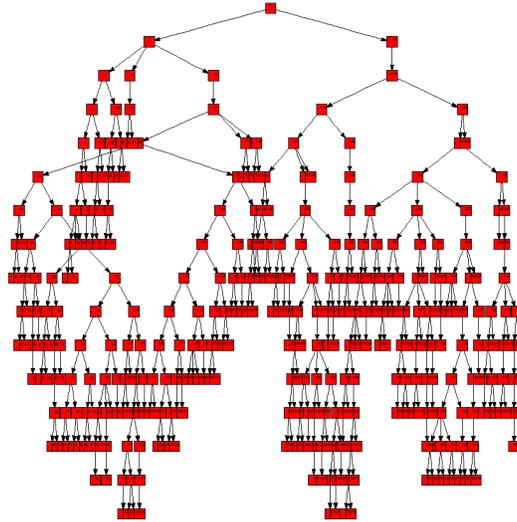


**Figura 3.** Curva de convergência, obtida a partir da mediana do EMQ obtido do total de 10 experimentos realizados.

É possível observar na imagem apresentada na Figura (4) a complexidade da solução gerada pelo processo evolutivo. Averiguou-se alta diversidade nos valores que caracterizam os nós terminais e não terminais. Também é factível analisar o tamanho da função gerada por meio do número de níveis da árvore, no caso da melhor solução global, possui 16 níveis e 388 nós para o grafo final. Para acessar a função gerada pela solução veja o link em: <https://1drv.ms/t/s!AiAix-tWooVygYI0Wxe5iLPSgtOZGA>

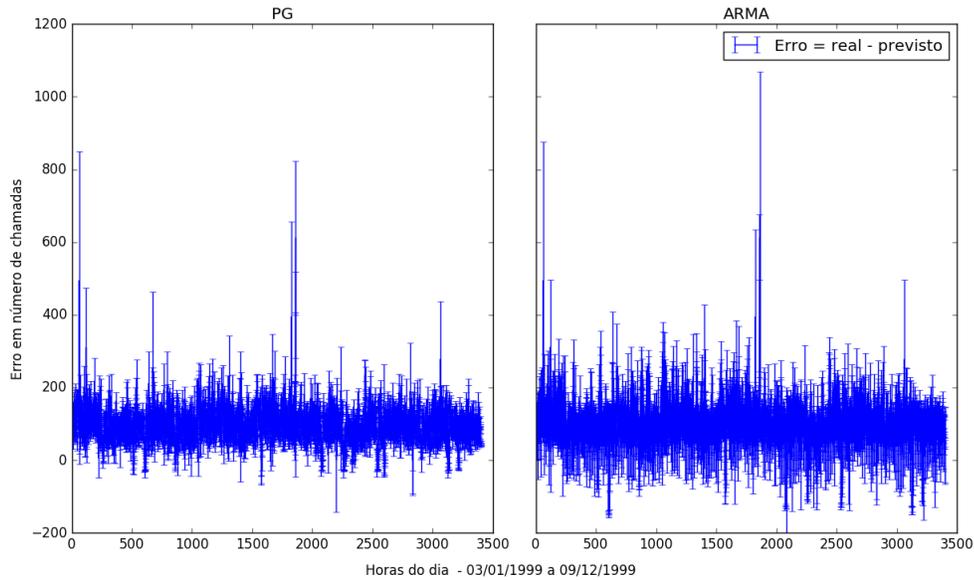
O gráfico da Figura (5) caracteriza um dos passos para averiguar a acurácia da previsão realizada pelos dois modelos. Tal passo consiste nas previsões geradas pela melhor solução (programa gerado nos experimentos) da PG e nos valores do ajuste do modelo ARMA em comparação com a base de dados real. Uma característica observada é a dificuldade de adaptação dos modelos aos pontos discrepantes (picos) encontrados na base de treinamento. Na Figura (5) o modelo da PG parece buscar uma maior adaptação aos pontos discrepantes.

Na Tabela 2 são apresentadas algumas estatísticas quanto ao processo de evolução/ajuste dos modelos. A tabela mostra o elevado valor do desvio padrão, fato que advém dos pontos discrepantes da série e da dificuldade dos modelos em tratar tal questão. Além da soma do quadrado dos resíduos (diferença entre ajuste e valor real ao quadrado), apresenta-se algumas estatísticas do RMEQ, como média e quartis. As estatísticas apresentam melhores indicadores para a PG quanto ao ajuste aos dados.



**Figura 4.** Árvore sintática, obtida a partir do melhor resultado entre todos os experimentos.

Já na Figura (6) é observado em detalhes por hora os dados reais e a previsão gerada pelo melhor indivíduo proveniente da PG e o resultado da previsão para o modelo ARMA. Observa-se que neste caso foram utilizadas chamadas realizadas pela central de teleatendimento em um intervalo de datas diferente do utilizado no treinamento/ajuste das funções. Na Tabela 3 apresenta-se os mesmos indicadores utilizados na evolução/ajuste. Observa-se que este período em questão, que envolve cerca de 140 horas de previsão, os indicadores para o modelo de PG são melhores do que o modelo ARMA.



**Figura 5.** Erros obtidos com os dados reais menos os valores gerados (processo de evolução) pelo modelo de programação genética e pelo modelo(ajuste) ARMA para os períodos de 03/01 a 09/12.

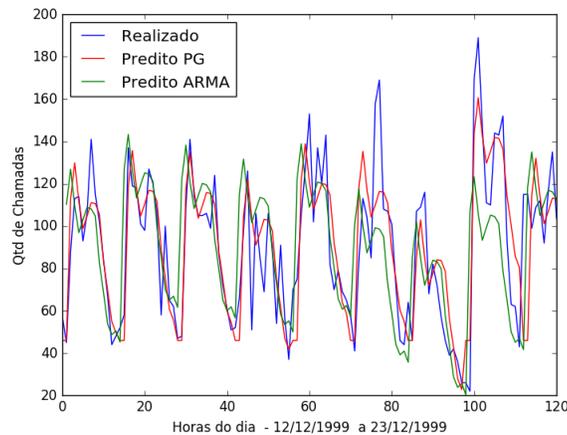
**Tabela 2.** Tabela comparativa com estatísticas para o período de evolução/ajuste dos dados - Somas dos quadrados dos resíduos e RMEQ considerando todas as horas do período de previsão de 03/01 a 09/12

	Modelo ARMA	Modelo PG melhor solução
Soma dos quadrados dos resíduos	6678236	<b>1883564</b>
<i>RMEQ</i> Média	44.30	<b>23.57</b>
<i>RMEQ</i> Desvio padrão	86.86	<b>54.66</b>
<i>RMEQ</i> Quartil 1	10.44	<b>5.30</b>
<i>RMEQ</i> Mediana	22.17	<b>11.81</b>
<i>RMEQ</i> Quartil 3	40.25	<b>21.03</b>

## 5 Conclusão e observações finais

Observa-se com a realização do estudo de caso que a abordagem da programação genética para previsão de chamadas de centrais de teleatendimento é válida. Resultado apresentado para PG é superior ao encontrado com a utilização de método largamente utilizado na literatura: ARMA com variáveis externas regressoras. Também é possível observar que a utilização da PG apresenta vantagens em relação aos outros modelos aplicados na previsão de chamadas.

Na Programação Genética não existem premissas para a validade do modelo e possibilita lidar com características lineares e não lineares dos dados, além de



**Figura 6.** Dados reais da base de dados e previstos pela função criada pela programação genética e o modelo ARMA para os períodos de 12/12 a 23/12.

**Tabela 3.** Tabela comparativa com estatísticas para o período de previsão - Somas dos quadrados dos resíduos e RMEQ considerando todas as horas do período de previsão de 12/12 a 23/12

	Modelo ARMA	Modelo PG melhor solução
Soma dos quadrados dos resíduos	167500	<b>49257</b>
<i>RMEQ Média</i>	29.11	<b>19.61</b>
<i>RMEQ Desvio padrão</i>	18.73	<b>13.33</b>
<i>RMEQ Quartil 1</i>	15.51	<b>4.26</b>
<i>RMEQ Mediana</i>	26.26	<b>10.97</b>
<i>RMEQ Quartil 3</i>	40.32	<b>21.51</b>

não haver necessidade de modelagem explícita da função. A grande desvantagem de empregar a PG é o alto tempo gasto em cada experimento, fazendo necessária a utilização de métodos paralelos de programação, ou máquinas com alto poder de processamento.

Acredita-se que seja interessante adicionar outras características da série na definição dos nós terminais, ou seja, processo de extração de características poderia tornar a representação da PG mais robusta. É necessário também testar diferentes métodos de desenvolvimento dos operadores genéticos e do processo de expansão das árvores ou mesmo inserir tais parâmetros no processo de evolução da solução.

O processo de treinamento não foi capaz de tratar completamente os pontos discrepantes apresentados na base de dados. Este ponto deve ser melhor explorado avaliando-se questões de sobre ajuste dos modelos. Por fim, seria interessante avaliar se uma hierarquia de modelos, com expressões específicas para o dia da semana, por exemplo, não poderia gerar melhores resultados.

## 6 Referências Bibliográficas

1. Alfaro Cid, E., Sharman, K., Esparcia-Alcazar, A.: Genetic programming and serial processing for time series classification 22 (09 2013)
2. Dasic, D., Kostic-Stankovic, M.: Contact centers organization, communication channels, and contact centers employees 44, 116–135 (01 2015)
3. Davidson, J.W., Savic, D.A., Walters, G.A.: Symbolic and numerical regression: Experiments and applications. *Inf. Sci. Inf. Comput. Sci.* 150(1-2), 95–117 (Mar 2003), [http://dx.doi.org/10.1016/S0020-0255\(02\)00371-7](http://dx.doi.org/10.1016/S0020-0255(02)00371-7)
4. Ding, S.: Workforce management in call centers - Forecasting, staffing and empirical studies. <https://goo.gl/HN7AJr> (2016)
5. Fortin, F.A., De Rainville, F.M., Gardner, M.A., Parizeau, M., Gagné, C.: DEAP: Evolutionary algorithms made easy. *Journal of Machine Learning Research* 13, 2171–2175 (jul 2012)
6. Graff, M., Escalante, H.J., Ornelas-Tellez, F., Tellez, E.S.: Time series forecasting with genetic programming. *Natural Computing* 16(1), 165–174 (Mar 2017), <http://dx.doi.org/10.1007/s11047-015-9536-z>
7. Higazi, S., Eddien Abdelhadi, W., Shalaby, R.: Application of genetic algorithm for the discovery of hidden temporal patterns in earthquakes data 9 (12 2013)
8. Ibrahim, R., L'Ecuyer, P.: Forecasting call center arrivals: Fixed-effects, mixed-effects, and bivariate models. *Manufacturing & Service Operations Management* 15(1), 72–85 (2013), <https://doi.org/10.1287/msom.1120.0405>
9. Ibrahim, R., Ye, H., L'Ecuyer, P., Shen, H.: Modeling and forecasting call center arrivals: A literature survey and a case study. *International Journal of Forecasting* 32(3), 865–874 (2016)
10. Lee, Y.S., Tong, L.L.: Forecasting time series using a methodology based on autoregressive integrated moving average and genetic programming. *Know.-Based Syst.* 24(1), 66–72 (Feb 2011), <http://dx.doi.org/10.1016/j.knosys.2010.07.006>
11. Mandelbaum, A., Sakov, A., Zeltyn, S.: Empirical Analysis of a call center. <http://iew3.technion.ac.il/serveng/References/ccdata.pdf> (2017)
12. Michael, S., Hod, L.: Symbolic Regression of Implicit Equations, pp. 73–85. Springer US, Boston, MA (2010), [http://dx.doi.org/10.1007/978-1-4419-1626-6\\_5](http://dx.doi.org/10.1007/978-1-4419-1626-6_5)
13. Morettin, P.A.: *Econometria Financeira: Um Curso em Series Temporais Financeiras*. Edgard Blucher (2011)
14. Poli, R., Langdon, W.B., McPhee, N.F.: A field guide to genetic programming. Published via <http://lulu.com> and freely available at <http://www.gp-field-guide.org.uk> (2008), (With contributions by J. R. Koza)
15. Riolo, R.L., Moore, J.H., Kotanchek, M.E. (eds.): Genetic Programming Theory and Practice XI [GPTP 2013, University of Michigan, Ann Arbor, May 9-11, 2013]. Springer (2014), <http://dblp.uni-trier.de/db/conf/gptp/gptp2013.html>
16. Sharp, D.E.: *Call Center handbook: Design, Operation and Maintenance*. Elsevier Science (2003)
17. Shen, H., Huang, J.Z.: Interday forecasting and intraday updating of call center arrivals. *Manufacturing & Service Operations Management* 10(3), 391–410 (2008), <http://EconPapers.repec.org/RePEc:inm:ormsom:v:10:y:2008:i:3:p:391-410>
18. Weinberg, J., Brown, L.D., Stroud, J.R.: Bayesian forecasting of an inhomogeneous poisson process with applications to call center data. *Journal of the American Statistical Association* 102(480), 1185–1198 (2012), <http://dx.doi.org/10.1198/016214506000001455>