



Modelagem de banco de dados para diagnóstico de transtornos mentais pela voz

Database modeling for diagnosis mental disorders by voice

Modelado de base de datos para el diagnóstico de los trastornos mentales de voz

Estefânia Mayumi Fuzy¹, André Armani², Sylvio Barbon Junior³

RESUMO

Descritores: Voz; Base de dados; Transtornos mentais; Diagnóstico

Objetivo: Desenvolver um modelo para o banco de dados de vozes brasileiro para armazenar dados do paciente e amostras de áudio capturados, mantendo um histórico de avaliação e tratamento, e possibilitar o diagnóstico automatizado de doenças mentais. **Métodos:** Pesquisar sobre testes relacionados ao diagnóstico tradicional e pela voz das doenças mentais com a finalidade de promover a abstração de entidades e atributos necessários na constituição da base de dados. Criação do Modelo Entidade-Relacionamento, mapeamento para o Modelo Relacional e criação do Diagrama de Classes. **Resultados:** Criação do modelo para o Banco de Dados de Vozes Brasileiro. **Conclusão:** A modelagem foi desenvolvida, de forma que a base de dados possa ser construída e utilizada.

ABSTRACT

Keywords: Voice; Database; Mental disorders; Diagnosis

Objective: Develop a model for the database of Brazilian voices to store data of patients and captured voices samples. This model supports the automated diagnosis of mental disorders. **Method:** Research about tests related to traditional diagnosis and based on voice of mental diseases in order to find the entities and attributes required for database. Creation of the entity-relationship model, mapping to a relational model and the class diagram. **Results:** Creation the model for the database of Brazilian voices. **Conclusion:** The modeling was developed and the database can be built and performed.

RESUMEN

Descriptores: Voz; Base de datos; Trastornos mentales; Diagnóstico

Objetivo: Desarrollar un modelo para la base de datos de voces brasileñas para almacenar los datos del paciente y muestras de audio captadas por mantener una historia de la evaluación y el tratamiento, y poder hacer un diagnóstico automatizado de la enfermedad mental. **Métodos:** Se realizó una búsqueda en los diagnósticos tradicionales conexos y la voz de lo trastorno mental con el fin de promover la captación de entidades y atributos requeridos en la constitución de las pruebas de base de datos. Creación del modelo Entidad-Relación, Modelo Relacional para el mapeo y la creación del diagrama de clases. **Resultados:** La creación del modelo de base de datos para las voces brasileñas. **Conclusión:** El modelo fue desarrollado para que la base de datos se puede construir y usada.

¹ Graduanda do curso Ciência da Computação, Universidade Estadual de Londrina - UEL, Londrina (PR), Brasil.

² Especialista em Cirurgia de Cabeça e Pescoço, Professor Assistente da Universidade Estadual de Londrina - UEL, Londrina (PR), Brasil.

³ Doutor em Física Aplicada Computacional, Professor Adjunto da Universidade Estadual de Londrina - UEL, Londrina (PR), Brasil.

INTRODUÇÃO

Um dos primeiros registros sobre doenças mentais que se tem conhecimento está na Bíblia, onde há descrição de patologias mentais semelhantes às doenças hoje diagnosticadas⁽¹⁾. Embora seja um problema antigo, inerente a espécie humana, apenas no século XX os recursos terapêuticos eficazes surgiram.

De acordo com pesquisa realizada pela Organização Mundial da Saúde, pelo menos 5% da população mundial sofre de depressão, um transtorno mental considerado comum, definido pela tristeza, perda de interesse, sentimento de culpa, distúrbios de sono ou apetite entre outros. Embora exista terapia para depressão, somente metade dos pacientes tem acesso ao tratamento médico. Em casos moderados a graves, é necessário que os pacientes sejam medicados e em seu estado mais grave, a depressão pode levar ao suicídio⁽²⁾.

Diagnosticar um paciente com depressão e avaliar o seu risco de suicídio para um psiquiatra é uma decisão muito importante, complexa e exigente⁽³⁾. Além do *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV)*, o diagnóstico baseia-se nos sintomas descritos, na história de vida e no comportamento do paciente durante a consulta.

Durante um discurso, a fala do indivíduo pode ser relacionada a sua condição emocional e mental, dessa forma, há como prever o nível de gravidade do transtorno afetivo do paciente, analisando a produção e a mediação nos resultados da fala⁽⁴⁾.

Propriedades acústicas da voz têm sido identificadas como indícios de depressão e existem evidências que determinados parâmetros da voz podem ser utilizados para diferenciar objetivamente entre a fala depressiva e a suicida⁽⁵⁾.

Ao avaliar a fala como um sinal digital de voz, pode-se extrair diversas informações, como a condição monótona que está relacionada com a ênfase reduzida. Essa avaliação pode ser medida por meio do nível da voz, ou quanto a inflexão, que é a variação da frequência fundamental⁽⁶⁾.

Apesar da existência de bancos de dados de vozes internacionais, como o KAY⁽⁷⁾, não existe algo parecido para o idioma português. O KAY dá suporte à análise acústica de vozes de pacientes que sofrem de algum transtorno para aplicações médicas ou de pesquisa. Este tipo de base inclui amostras de pacientes com uma ampla variedade de distúrbios da voz, sejam eles orgânicos, neurológicos, traumáticos ou psicogênicos. É sabido que a fala está relacionada a avaliação da saúde mental pela voz⁽⁴⁻⁵⁾, logo, nesta avaliação, o idioma é de grande importância, e com a base no idioma português é possível realizar soluções automatizadas e de apoio a decisão médica.

O uso de Sistemas de Informação possibilita a criação de soluções que facilitem e ajudem médicos e pesquisadores, como o diagnóstico automatizado de doenças que não tem o objetivo de se tornar absoluto, de forma a descartar a análise médica, mas sim apoiar a decisão médica. Al-Haj et al.⁽⁸⁾ propuseram uma

metodologia para segmentar automaticamente, glândulas de imagens histopatológicas do tecido glandular da próstata para classificação de grau, que descreve a anormalidade da célula cancerígena e seu grau de agressividade através de uma escala numérica, com resultados entre 91% e 97% de precisão, sensibilidade e especificidade. Para o diagnóstico automatizado do Alzheimer, outro estudo adotou a espessura de regiões do córtex cerebral em imagens de ressonância magnética como características para classificação, baseado em pesquisas sobre neuropatologia⁽⁹⁾.

A utilização do diagnóstico por meio da voz, possibilita a avaliação da condição do paciente de regiões remotas, que muitas vezes não tem fácil acesso a esse tipo de profissional. A voz pode ser armazenada, não é necessário que um especialista da área médica adquira a voz, e posteriormente avaliada pelo especialista. Por exemplo, um usuário do sistema de informação enviado a uma comunidade no interior do país faria coleta das vozes, o Sistema de Informação efetuaria o pré-diagnóstico e triagem. Os pacientes com os resultados que apontem atenção seriam encaminhados para o especialista da área da saúde.

Com o uso da base de dados, é possível acompanhar a evolução do quadro do paciente de forma rápida e organizada, com base nas consultas armazenadas, para a geração de um histórico e quadro de progressão.

Considerando base de dados em saúde, tem-se atualmente a preocupação com o volume de informação persistida. As aplicações *big data* podem ser utilizadas no domínio de serviço do *e-health* na redução de custos e melhoria de desempenho contribuindo no acompanhamento do paciente⁽¹⁰⁾. O *e-health* tem como núcleo o uso de serviços ou sistemas de tecnologia da informação em favor da saúde⁽¹¹⁾, que baseado em *big data* prove aplicações que lidam com coleções complexas e volumosas de informações como imagem, vídeo e texto. O volume de informações geradas pelo *e-health* alimentadas por prontuários eletrônicos e aplicações de monitoramento é cada vez maior. Há trabalhos que propõem o *BDeHS (Big Data e-health Service)* com o objetivo de executar aplicações *big data* no domínio de serviços *e-health*, que fornece gerenciamento de dados para melhorar o processo de *stream*, gerenciamento operacional e conformidade regulatória⁽¹⁰⁾. Desta forma é possível alcançar o processamento em tempo real, reconfiguração e aprimoramento da capacidade da rede, acrescentando qualidade e velocidade aos processos médicos.

O objetivo desse trabalho é modelar uma estrutura para prover o armazenamento e recuperação de informações para o diagnóstico tradicional ou automatizado de doenças mentais pela voz.

DIAGNÓSTICO DE DOENÇAS MENTAIS PELA VOZ

No trabalho, foram realizados estudos com amostras de áudio da voz de homens e mulheres saudáveis e com diagnóstico de depressão e alto risco de suicídio a partir de um banco de dados de voz⁽⁵⁾. Seu objetivo era investigar

as propriedades acústicas da fala como um sintoma psicomotor da depressão e do risco de suicídio, e para testar sua relevância como um parâmetro diferencial para o diagnóstico. O estudo foi separado em duas partes, a primeira analisa características vocais de mulheres depressivas não medicadas e a segunda, a fala de homens depressivos e com alto risco de suicídio. Segundo o autor, a maior desvantagem dessa abordagem foi não ter controle sobre as especificações técnicas do equipamento de aquisição do áudio, ambiente e procedimentos. Devido a qualidade acústica associada com ruídos do ambiente, uso de equipamentos de gravação de baixa qualidade e a utilização de procedimentos não padronizados o trabalho pode ter resultados insatisfatórios. Isso se deve as características das amostras, pois eram gravações realizadas durante as consultas e conversas telefônicas entre o paciente e o médico. Foi concluído, que as características vocais do comportamento do formante e o *power distribution* podem ser utilizadas para distinguir indivíduos que carregam o diagnóstico de depressão e alto risco de suicídio.

Discursos foram analisados para verificar alterações nas características acústicas causadas pela depressão⁽⁴⁾. Foi relatado que a fala pode estar relacionada com a condição emocional e mental enquanto o locutor está falando, além de ajudar a prever o nível de gravidade do transtorno, pois a patologia afeta a produção e mediação dos resultados da fala. O banco de dados utilizado continha gravações de pacientes categorizados como depressivos ou saudáveis por um médico externo ao projeto. As amostras foram gravadas em mono canal, digitalizadas com a taxa de amostragem de 16 kHz e convertidas para 10 kHz. Os sinais amostrados foram editados para remover a voz do entrevistador, pausas maiores que 0.5 segundos e ruídos de fundo, como espirros e movimentação de portas. Os resultados experimentais mostraram que com o cálculo da entropia é possível separar as classes com altas porcentagens de validações corretas. Sendo a entropia definida como a quantidade média de informação significativa que a fala leva ao ouvinte.

A diferença entre a fala livre e a fala automática foi abordada por estudiosos, pois a primeira requer atividades cognitivas, como procurar por uma palavra e planejar o discurso, além da atividade motora. Foram analisados pacientes idosos, separados em dois grupos com características diferentes de depressão⁽¹²⁾. Os testes foram baseados na contagem de 1 a 10 de forma crescente e depois decrescente. Estas atividades tem justificativa na exigência do processo cognitivo. Para a avaliação da fala livre, os pacientes foram envolvidos em uma entrevista, com tópicos de conversação como: atividades recentes e programas de TV favoritos. Outros tópicos foram relacionados ao relato durante 5 minutos de experiências positivas e negativas. A partir dessas atividades, foi concluído que as impressões clínicas são substancialmente relacionadas aos parâmetros acústicos, que as variações temporais refletem o estado depressivo, enquanto as características prosódicas refletem o traço depressivo.

Devido a necessidade de ter uma base de dados de

distúrbios da voz, de acordo com as condições patológicas de pacientes e características do mandarim, Wanget al, desenvolvam uma abordagem composta pela sustentação de uma vogal, locução de palavras, sentenças e poesia⁽¹³⁾. A primeira etapa contém as vogais mais utilizadas, a segunda, dez palavras simples, a terceira, duas sentenças com cinco palavras e uma de sete, por fim, a última etapa exige a locução de uma poesia popular antiga. O banco de dados fornece o armazenamento e recuperação de uma escala que avalia a percepção vocal no nível glótico, para cada amostra, avaliadas por otorrinolaringologistas. Além de informações quanto ao gênero, idade, conclusão do diagnóstico e exemplos padrões para comparação entre o mandarim e o inglês.

MÉTODOS

A modelagem da base de dados foi realizada avaliando os dados do paciente que estão presentes em testes psicológicos e quais os parâmetros relacionados a sua fala são extraídos e analisados segundo o estado da arte para o diagnóstico de doenças mentais. O mecanismo de obtenção da voz por meio das técnicas de processamento digital de sinais também foi estudado para que as informações armazenadas propiciem a aplicação de quaisquer técnicas de extração de características da voz.

Duas premissas foram utilizadas como guia na confecção da modelagem, a primeira: quais os tipos e tamanhos mais adequados para o objetivo, de modo que não haja perda de qualidade do dado e desperdício de memória, para que não influencie negativamente no diagnóstico automatizado. Por exemplo, caso tenha sido escolhido um formato de armazenamento de voz que ocupe pouca memória, o sinal pode se tornar de baixa qualidade e tal informação fica presente na base para auxiliar na escolha das abordagens de processamento. A segunda, foi a preservação de informações relevantes na anamnese do paciente, o que dá suporte a preservação de dados para os testes tradicionais.

Na fase de criação da modelagem, o primeiro passo foi elaborar o Modelo Entidade-Relacionamento (MER), um modelo de dados utilizado para o “projeto conceitual de aplicações de banco de dados”⁽¹⁴⁾, nele são descritos entidades, relacionamentos e cardinalidade. A partir do esquema conceitual, foi projetado um esquema de banco de dados relacional, para isso foi utilizado o algoritmo de mapeamento do MER para o Modelo Relacional (MR), que representa o banco de dados como um conjunto de relações e possui forte fundamentação matemática. Esse passo é necessário para que sejam respeitadas as restrições de integridade do modelo, restrição de domínio, restrição de chave, integridade de entidade e integridade referencial, de forma que não aconteçam atualizações errôneas no banco⁽¹⁴⁾. Então, foi criado o Diagrama de Classes, muito semelhante ao MER, em que são exibidas as entidades a serem criadas com seus respectivos atributos e tipos, seus relacionamentos e cardinalidades, além de especificar a estrutura do esquema do banco de dados⁽¹⁴⁻¹⁵⁾.

Uma entidade é algo no mundo real que pode ser

abstraido como uma modelo, como uma pessoa, que é um objeto que existe fisicamente, ou um curso universitário, que é conceitual⁽¹⁴⁻¹⁵⁾. Cada entidade possui atributos que a descrevem, por exemplo, ao criar uma entidade *Pessoa*, ela pode ter os atributos *nomePessoa*, *cpf*, *telefone*, *dataNascimento* e *idade* e uma entidade *CursoUniversitario* com o *nomeCurso* e *dataCriação*.

Os atributos podem ser classificados como simples (atômicos) ou compostos. O primeiro seria um valor não divisível, o campo *cpf* com o valor “934.282.283-59” por exemplo, e o segundo pode ser dividido em partes menores independentes e é formado pela concatenação dessas partes, tal como o atributo *nomePessoa*, que pode ser dividido em *primeiroNome* e *ultimoNome*. Eles podem ser classificados como únicos, de forma que só existe um para uma pessoa ou multivalorado, que pode ter um ou mais, como o *cpf* e o *telefone*, respectivamente. Podem ser atributos armazenados ou derivados, a *dataNascimento* da pessoa por exemplo, é armazenado e a *idade* é derivado da *dataNascimento*, pois se essa informação for armazenada, será necessário alterá-la todo ano. E quando não tem valores aplicáveis a ele, o atributo pode ser *null*, como quando o número do telefone não é conhecido, o atributo ficará vazio⁽¹⁴⁾.

Em entidades denominadas fortes, existe um atributo, ou mais, que pode identificá-la unicamente, chamado atributo-chave⁽¹⁴⁾. Estes precisam ser únicos no conjunto e quando mais de um, a sua combinação se torna uma chave primária, que não pode ser nulo, pois ele será utilizado para identificação, como o atributo *cpf* da *Pessoa*. Uma entidade é dita fraca, quando ela não tem o atributo-chave e precisa ser identificada pela entidade com quem ela se relaciona, dessa forma, ao realizar o mapeamento, o(s) atributo(s) da chave primária da entidade identificadora são adicionados à entidade fraca como chave estrangeira⁽¹⁴⁾.

As entidades podem se relacionar entre si, como entre *Pessoa* e *CursoUniversitario* pode existir o relacionamento *Estuda*. Um tipo de relacionamento pode ter diferentes graus, dependendo do número de entidades participantes, se forem duas será grau dois, também denominado binário, se forem três será grau três, ou ternário e assim por diante. Tipos de relacionamentos binários podem ter as cardinalidades: a) 1:1: uma pessoa pode se relacionar com apenas um curso; b) 1:N: uma pessoa pode se relacionar a vários cursos; c) N:1: várias pessoas podem se relacionar com apenas um curso; d) M:N: várias pessoas podem se relacionar com vários cursos⁽¹⁴⁾.

A agregação permite “a criação de objetos compostos com base em seus objetos componentes”. Ela pode ser utilizada em três casos, primeiro quando é necessário agregar valores de atributos para formar um objeto, segundo quando é necessário representar um relacionamento de agregação como comum e terceiro quando o objeto o objeto de agregação precisa se relacionar com outro⁽¹⁴⁾.

A implementação de um protótipo do modelo criado foi realizado com o auxílio da ferramenta PostgreSQL, devido a sua gratuidade e por ser *open source*, o que facilita as implementações específicas para o problema e sem

nenhum custo. O PostgreSQL é um sistema de banco de dados objeto-relacional *open source*⁽¹⁶⁾. Possui mais de quinze anos de desenvolvimento ativo e uma arquitetura comprovada que ganhou uma forte reputação de confiabilidade, integridade de dados e correção. Pode ser executado na maioria dos sistemas operacionais, apresenta total suporte para chave estrangeira, junção, visões, gatilhos e procedimentos de armazenamento. Além de tipos de dados como *Integer*, *Numeric*, *Boolean*, também suporta o armazenamento de grandes objetos binários, incluindo fotos, sons e vídeos.

O modelo desenvolvido terá uma ampla divulgação e poderá ser encontrado e obtido no site do grupo de pesquisa.

RESULTADOS E DISCUSSÃO

Após a criação do MER e feito seu MR, foi criado o Diagrama de Classes, como mostra a Figura 1. Nas próximas subseções serão descritas as entidades, atributos, os relacionamentos e as duas cardinalidades.

Entidades

As entidades *Paciente* e *Medico* existem para representar as informações do paciente e do médico, para sua identificação. Outra entidade criada é a *Consulta*, mapeada como uma agregação, pois é a partir dela que o áudio poderá ser adquirido, em que ficarão armazenadas as informações da consulta e do paciente na sessão de amostragem da voz.

As entidades *Nacionalidade*, *Religião* e *Profissão* são todas informações acessórias sobre o paciente criadas para manter a normalização do banco de dados. Como o diagnóstico da doença será baseado na fala do paciente, a entidade *Nacionalidade* é importante, caso o paciente tenha outra língua materna⁽¹⁷⁾, o que pode influenciar no discurso.

Para armazenar os sinais de voz capturados durante a consulta, tem-se as entidades *Audio* e *FormatoArquivo* para armazenar qual o tipo do arquivo será gerado com base no sinal. A entidade *Audio* foi mapeada como uma entidade fraca, pois ela só existe se existir pelo menos uma consulta e um teste para se relacionar. Dessa forma, ao realizar o mapeamento para Entidade Relacional, as chaves primárias de *Consulta* e *Teste* são adicionadas como chave estrangeira e a sua combinação se torna chave primária do *Audio*.

A entidade *Teste* é o modelo dos testes para avaliação da fala do paciente, na entidade *Parametros* são armazenados os parâmetros necessários para a execução dos testes e a entidade *ParametrosTeste* vincula os parâmetros com seus respectivos testes. A última foi criada devido ao mapeamento do relacionamento com cardinalidade [N:N], pois não é possível representar esse tipo de relacionamento por uma única chave estrangeira de uma das entidades participantes⁽¹⁴⁾.

A entidade *DadosExtraídos* representa as informações extraídas a partir do sinal de áudio por meio de técnicas de processamento digital de sinais. A entidade *DadosExtraídos* é uma entidade fraca, pois sua existência

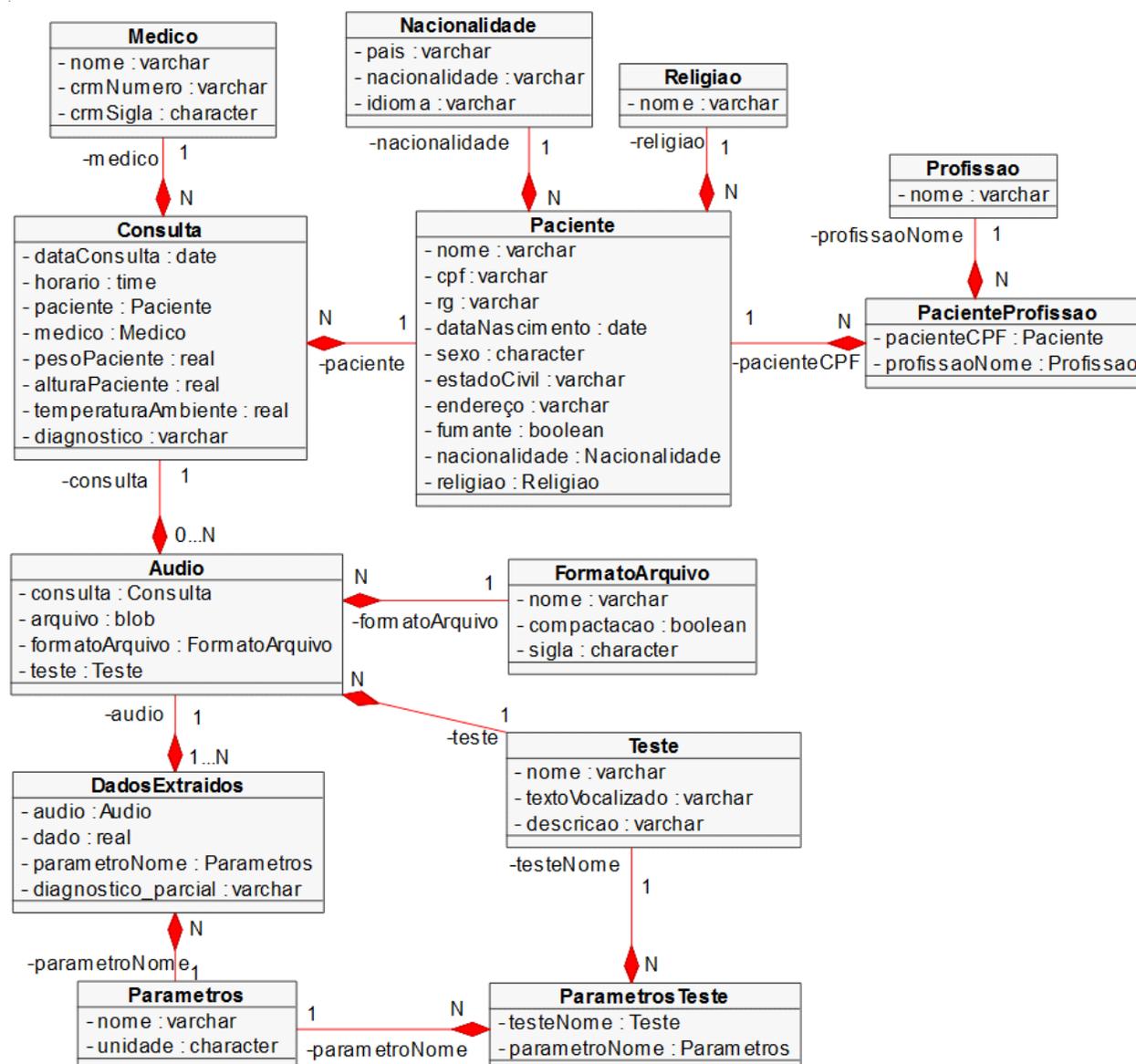


Figura 1 - Diagrama de classes da base de dados

depende da entidade *Audio* e da entidade *Parametros*, e pelo mapeamento de uma entidade fraca, as chaves primárias das entidades identificadoras são adicionadas como chave estrangeira e em conjunto, se tornam chave primária. O processamento e o armazenamento correto desses dados são de grande importância, visto que eles serão necessários para a obtenção do diagnóstico automatizado.

Atributos

Na entidade *Medico*, existem os atributos *nome*, *crmNumero* e *crmSigla*, com o objetivo de armazenar informações sobre o médico, e identificá-lo unicamente por meio dos atributos obtidos do registro do Conselho Nacional de Medicina (CRM).

Para o paciente, são necessários os atributos: *nome*, *cpf*, *rg*, *dataNascimento*, *sexo*, *estado Civil* e *endereco*. O fato do paciente fumar é relevante, pois segundo⁽¹⁸⁾ tal fato tem influência sobre os parâmetros acústicos da voz, seja de modo isolado ou em conjunto com o etilismo, tosse, hábito de pigarrear, entre outros.

Para identificar uma consulta de forma única, são necessários os atributos *dataConsulta*, *horário*, além do *paciente* e o *médico*. O *peso* e a *altura* do paciente são importantes, pois um Índice de Massa Corporal alto ou baixo pode estar diretamente relacionado a depressão, porém de formas distintas entre os gêneros masculino e feminino, em que, por exemplo, quanto mais alto o IMC nos homens, pode resultar em uma forma mais severa da depressão, enquanto nas mulheres a relação pode não ser a mesma, além de que o ganho ou a perda de peso serem sintomas da doença⁽¹⁹⁻²⁰⁾. Tem-se também o atributo *diagnostico*, que armazenará qual a doença mais provável do paciente, com base nos diagnósticos parciais dos testes executados e/ou na decisão do médico (depende se o diagnóstico será automatizado ou não), com o objetivo de manter o histórico do paciente.

O sinal de áudio capturado durante a consulta está representado na entidade *Audio*, no atributo *arquivo*, como um tipo de dado binário. Os demais atributos foram incluídos por meio do mapeamento da entidade, sendo o atributo *consulta*, um identificador da consulta que gerou

esse áudio, o campo *formato.Arquivo* determina qual será o tipo desse do sinal de áudio e o *teste* diz para qual teste essa gravação será utilizada.

Na entidade *Formato.Audio*, tem-se os atributos *nome*, *compactação* e *sigla*. O primeiro campo, deve ser preenchido com o nome formato do sinal de áudio, como “WAVEform Audio Format”, o segundo pode ser verdadeiro ou falso, de forma que indique se o sinal de áudio foi compactado ou não, e o terceiro preenchido com a sigla do formato, como “wav”, para o exemplo dado.

Para a entidade *Teste*, são necessários o nome do teste para sua identificação, qual o texto vocalizado e a sua descrição, como por exemplo “Contar de 1 a 10”, “1, 2, 3, 4, 5, 6, 7, 8, 9, 10”, “Contar de forma livre, sem ler, os números no intervalo de 1 a 10”, respectivamente⁽¹²⁾. Cada teste é relacionado com diversos parâmetros, que podem ser encontrados na entidade *Parametros*, por meio dos atributos *nome* e *unidade*, que podem ser “frequência fundamental” e “Hz”, por exemplo.

Na entidade *DadosExtraídos*, há o atributo *audio*, que mostra a qual sinal de áudio o dado extraído pertence, o *parametroNome*, que relaciona o dado ao seu determinado parâmetro e o campo *dado*, que armazena a informação extraída. E o atributo *diagnóstico_parcial* para armazenar o resultado de cada teste.

Os atributos do tipo *varchar* têm, em sua maioria, a finalidade de armazenar informações de sequência de caracteres para identificação da entidade, como o campo *nome* de *Médico* e *Paciente* ou conteúdo como o campo *textoVocalizado* da entidade *Teste*. Os atributos *Medico.crmNumero* e *Paciente.cpf* são aparentemente numéricos, porém, como não é necessário realizar cálculos com esses valores, eles foram modelados como *varchar*.

O tipo *character* é utilizado para atributos de texto em que é possível delimitar o número de caracteres, por isso ele é utilizado nos atributos *Medico.crmSigla*, *Paciente.sexo*, *Formato.Arquivo.sigla* e *Parametros.unidade*, dessa forma, permite os valores ‘F’ e ‘M’ para o atributo *sexo*, sem que seja desperdiçado espaço se fosse utilizado o tipo *varchar*.

Para os atributos que devem aceitar números com casas decimais, como *Consulta.pesoPaciente* e *DadosExtraídos.dado* por exemplo, foi escolhido o tipo *real* por suportar até seis casas decimais, e não o *numeric*, pois não será necessária a realização de cálculo com esses valores.

Os atributos *Paciente.fumante* e *Formato.Arquivo.compactacao* são do tipo *boolean* pois as informações necessárias são, apenas, se o paciente fuma ou não, e se o sinal de áudio está compactado ou não.

Relacionamento

Para que uma consulta exista, precisa existir uma entidade médico e paciente vinculados a ela e a cardinalidade é [N,1] pois o médico/paciente pode ter várias consultas. Porém uma consulta pertence a apenas uma combinação de médico, paciente, data e hora. Para cada consulta podem ser capturados um ou mais amostras de áudio, por isso sua cardinalidade de [1,N], no entanto, o áudio só existe se estiver vinculado a uma consulta e a um teste.

Da mesma forma ocorre entre *Paciente-Nacionalidade* e *Paciente-Religião*. Em relação ao relacionamento entre *Paciente*

e *Profissao*, existe uma sub-entidade devido ao fato que um paciente pode ter mais de uma profissão e que a profissão pode ser vinculada com mais de um paciente⁽¹³⁾.

Cada sinal de áudio armazenado pode ter um formato de armazenamento, seja .wav ou .mp3 por exemplo. Logo, ele deve ser relacionado com a entidade *Formato.Arquivo* para que possa ser identificado em um relacionamento [N,1] pois cada sinal de áudio só tem um formato de áudio, e o formato de áudio pode pertencer a mais de um sinal. Ele se relaciona com *Teste*, pois para cada sinal de áudio, pode ser aplicado um teste, portanto, sua cardinalidade é [N,1].

Para cada teste é possível haver vários parâmetros relacionados, tal como vários parâmetros pertencem a vários testes. Dessa forma, foi criada a entidade *ParâmetrosTeste* que faz a ligação de testes e parâmetros. As entidades *Teste* e *ParâmetrosTeste* tem cardinalidade [1,N] pois cada teste pode ter vários parâmetros e, *Parâmetros* e *ParâmetrosTeste* tem cardinalidade [1,N] também, porque cada parâmetro pode pertencer a vários testes.

Após gravado o sinal áudio para determinado teste durante determinada consulta, e identificados quais os parâmetros o teste necessita para chegar a uma conclusão, eles podem ser extraídos do áudio original, por meio de técnicas de processamento digital de sinais específicas e armazenados conforme a entidade *DadosExtraídos*. Um registro nesta entidade só existe caso seja necessário um áudio, caso contrário não há como identificar sua origem e função. Com a cardinalidade [1,N] de *Audio* para *DadosExtraídos* podem ser extraídos vários dados de um áudio, e [N,1] de *DadosExtraídos* para *Parametros*, visto que cada dado pode pertencer apenas a um parâmetro, e que o mesmo parâmetro pode ser relacionado a dados distintos.

CONCLUSÃO

O objetivo de modelar o banco de dados foi atingido, portanto, ele poderá ser construído para a utilização em diferentes sistemas que necessitem armazenar e recuperar sinais da fala, ou que realizem o diagnóstico por meio dela. Para facilitar o entendimento e o desenvolvimento da base de dados, os modelos serão disponibilizados no site do grupo de pesquisa para livre acesso.

Para os trabalhos futuros serão divulgados o prontuário eletrônico com o uso da base modelada e processos de pré-diagnóstico. Além da aplicação de técnicas de banco de dados multimídia, tal como reconhecimento de voz para auxiliar no processo de aprendizado de máquina durante o desenvolvimento do diagnóstico automatizado e indexação baseada em texto de dados de áudio.

AGRADECIMENTOS

Agradecimentos ao PPSUS pelo financiamento de equipamentos contemplado pelo edital 04/2013 Programa Pesquisa para o Sistema Único de Saúde: Gestão Compartilhada em Saúde PPSUS Edição 2012, cujo protocolo é 41621.433.41669.12092013.

REFERÊNCIAS

- Varella D. Saúde Mental [Internet]. [citado 2014 Mar 19]. Disponível em: <http://drauziovarella.com.br/letras/s/saude-mental/>
- Varella D. Mais de 350 milhões de pessoas sofrem de depressão no mundo [Internet]. [citado 2014 Mar 19]. Disponível em: <http://drauziovarella.com.br/noticias/mais-de-350-milhoes-de-pessoas-sofrem-de-depressao-no-mundo/>
- Ozdas A, Shiavi R, Silverman SE, Silverman MK, Wilkes DM. Investigation of vocal jitter and glottal flow spectrum as possible cues for depression and near-term suicidal risk. *IEEE Trans Biomed Eng.* 2004;51(9):1530-40.
- Vayada M, Nayak V, Agrawal A. Analysis of speech of depressed person. *Sarjan.* 2013;1(2):33-6.
- France DJ, Shiavi RG, Silverman S, Silverman M, Wilkes DM. Acoustical properties of speech as indicators of depression and suicidal risk. *IEEE Trans Biomed Eng.* 2000; 47(7):829-37.
- Kraepelin E. Manic-depressive insanity and paranoia. London: E&S Livingstone; 1921.
- KayPentax. Disordered voice database and program, model 4337 [Internet]. [citado 2014 Mar 19]. Disponível em: [http://www.kaypentax.com/index.php?option=com_product&Itemid=3&controller=product&task=learn_more&cid\[\]=52](http://www.kaypentax.com/index.php?option=com_product&Itemid=3&controller=product&task=learn_more&cid[]=52)
- Al-Haj Saleh SN, Al-Kadi OS, Al-Zoubi MB. Histopathological prostate tissue glands segmentation for automated diagnosis. In: *Proceeding of The Second IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies*; 2013 Dec 3-5; The University of Jordan, Amman, Jordan.
- Huang L, Pan Z, Lu H. Automated diagnosis of Alzheimer's disease with Degenerate SVM-Based Adaboost. In: *Proceedings of the 5th International Conference on Intelligent Human-Machine Systems and Cybernetics*; 2013 Ago 26-27; Zhejiang University, Hangzhou, China.
- Liu W, Park EK. Big data as an e-health service. In: *Proceedings of the International Conference on Computing, Networking and Communications*; 2014 Fev 3-6; Honolulu, Hawaii, USA.
- Álvarez MM. International e-health and national health care systems. *Encyclopedia of Health Economics.* London: Elsevier; 2014.
- Alpert M, Pouget ER, Silva RR. Reflections of depression in acoustic measures of the patient's speech. *J Affect Disord.* 2001;66(1):59-69.
- Wang D, Zhang J, Fu Q, Yu P, Yang L, Yan Y, et al. An introduction to Mandarin disordered voice database. In: *Proceedings of the International Conference Audio Language Image Processing*; 2008 Jul 7-9; Shanghai, China.
- Elmasri R, Navathe B. *Sistemas de banco de dados.* 6a ed. São Paulo: Pearson Addison Wesley; 2011.
- Pressman RS. *Engenharia De Software.* 6a ed. Rio de Janeiro: McGrawHill; 2006.
- PostgreSQL [Internet]. [citado 2014 Mar 27]. Disponível em: <http://www.postgresql.org/>
- Sadock BJ, Sadock VA. *Manual de psiquiatria clínica: referência rápida.* 5a ed. Porto Alegre: Artmed; 2012.
- Pinto AGL, Crespo AN, Mourão LF. Influence of smoking isolated and associated to multifactorial aspects in vocal acoustic parameters. *Braz J Otorhinolaryngol.* 2014;80(1):60-7.
- Dragan A, Akhtar-Danesh N. Relation between body mass index and depression: a structural equation modeling approach. *BMC Med Res Methodol.* 2007;7:17.
- De Wit LM, van Straten A, van Herten M, Penninx BW, Cuijpers P. Depression and body mass index, a u-shaped association. *BMC Public Health.* 2009;9:14.