

Como a extração de personalidade através do teclado pode beneficiar a personalização na Educação

Sandy M. Porto, Wanderson S. Costa, Maria Augusta S. N. Nunes, Leonardo N. Matos

Departamento de Computação
Universidade Federal de Sergipe (UFS) – São Cristóvão, SE – Brasil, CEP 49100-000

sandyporto@gmail.com, wandersonscosta@hotmail.com,
gutanunes@dcomp.ufs.br, lnmatos@ufs.br

Abstract. *This paper discusses how you can extract some personality traits, based on The Big Five Model through typing rhythm. Also, discusses how the extraction of some personality traits of the student might facilitate the customization of educational content.*

Resumo. *Este artigo discute como é possível extrair alguns traços da personalidade, baseado no The Big Five Model, através do ritmo de digitação. E ainda, discute como a obtenção dos traços de personalidade do aluno pode facilitar a personalização do conteúdo educacional.*

1. Introdução

O Marketing sempre valorizou a identificação do comportamento social e a interação entre as pessoas como algo a ser explorado. Saber identificar quais são as ambições ou necessidades do cliente (ou um grupo deles) e, assim, personalizar ofertas a ele, pode significar, além da satisfação do cliente, também, altas margens de lucro. Na Educação, estratégias de Marketing também devem ser aplicadas, visto que o aluno (ou o conjunto deles) nada mais é, num primeiro momento, que um consumidor, que deve ser seduzido, atendido e motivado. Saber onde as informações de interesse do cliente estão, qual caminho a percorrer, ou onde iniciar a coleta e/ou disseminação dessa informação para que estas cheguem finalmente ao seu cliente, é fundamental.

Na internet o que encontra-se com frequência são programas que propõem a interação social, ou que tenham como objetivo agrupar pessoas por afinidades. A interação social pode ser obtida através de programas de computador intitulados vulgarmente como *chats online*, que nada mais são que programas de conversação. Como exemplo desses programas tem-se: *Yahoo Messenger*, o *ICQ*, *Google Talk*, *Windows Live Messenger*, esse último o mais utilizado atualmente.

Hoje, na internet, principalmente entre os jovens, a grande febre é o uso das redes sociais, que promovem o encontro virtual de pessoas com afinidades em comum, como por exemplo: pelo gosto musical, vide *Last.fm*, *MySpace*; afinidades conjugais, *Badoo*, *ParPerfeito*; afinidades profissionais, *LinkedIn*, e poderiam ser citados vários outros, inclusive redes sociais que promovem encontros de pessoas com afinidades políticas, educacionais, cinematográficas, até mesmo por proximidade de moradia. Estes são os nichos continuamente potencializados por novas redes sociais. Atualmente as grandes redes sociais são o *Orkut*, *Twitter* e *Facebook*.

Todos os programas que suportam as redes sociais citadas anteriormente têm uma tendência, ou mesmo uma necessidade de prever e/ou construir um mapeamento psicológico dos seus usuários, seja para fins de manutenção da própria “vida útil” do programa, objetivando promover mais interação entre os usuários ou mesmo descobrir a qual tipo de informação esses usuários gostariam de ter acesso; para fins de mercado, como vender espaços publicitários, promover aplicativos que explorem essas necessidades e interesses pessoais dos usuários e finalmente, também, para fins de educação, onde a necessidade de personalização de ambientes educacionais e conseqüente adaptação de estratégias de ensino de acordo com cada modelo de aluno se torna fundamental [Self 1999].

Nas redes sociais esse interesse do usuário pode ser medido através das pessoas com quem ele se relaciona, e por qual tipo de conteúdo ele se interessa. Nos programas de conversação isso pode ser extraído através do contexto das mensagens escritas, através de *text-minning*, ou palavras chaves.

Obtendo essas informações sobre o usuário é possível sugerir algo a esta pessoa, como se sugere algo a um amigo ou familiar. Note que conhecendo bem o usuário é possível prever o que é de interesse (ou não) do usuário, esse processo de personalizar intitula-se de Sistemas de Recomendação. A recomendação pode ser de um produto, um texto, um site, uma sugestão de amizade, uma sugestão de contato profissional, um objeto educacional, um grupo com interesses similares (descrito em mais detalhes na seção 4).

Há estudos [Hu e Pu 2009] que afirmam que Sistemas de Recomendação baseados em personalidade têm atraído a atenção de pesquisadores e profissionais da indústria. Pesquisas têm mostrado que personalidade é um fator primário e duradouro que determina o comportamento humano e que há uma significativa conexão entre personalidade, os gostos e interesses das pessoas. É, portanto, razoável acreditar que Sistemas de Recomendação baseados em personalidade possam prover mais informações/serviços personalizados, desde que eles possam entender melhor os consumidores da perspectiva psicológica [Nunes e Cazella 2011].

No entanto, em [Brinkman e Fine 2008] percebe-se que personalidade é comumente obtida pela aplicação de questionários com mais de 100 questões, o que requer algum tempo e esforço por parte do usuário para respondê-los e exigir isso do usuário não é uma abordagem agradável. Ainda em [Brinkman e Fine 2008] nota-se que continuamente surgem pesquisas que tentam obter personalidade de uma forma menos intrusiva, por exemplo, analisando gravações automáticas de dados da interação do usuário com o sistema. Um desses dados de interação que podem ser facilmente gravados automaticamente sem ser intrusivo é o ritmo de digitação e o uso do mouse.

Existem muitos trabalhos descritos em [Nunes 2011] que tentam extrair a personalidade dos usuários a partir de técnicas não intrusivas, em [Brinkman e Fine 2008], por exemplo, eles avaliam a escolha dos usuários para customização de algumas aplicações, a saber, *Windows Media Player*, *ICQPlus* e o *Winamp*, através da tecnologia de *skins*, e dependendo da escolha, pode-se prever a personalidade do usuário, fazendo comparações com a escolha de outros usuários.

2. Personalidade

A personalidade representa a estrutura e a dinâmica de um conjunto de características de um indivíduo normalmente adquirido do ambiente e das experiências pessoais. É mais do que apenas aparência física e é relativamente estável. No entanto, é necessário que se possa apresentar uma personalidade e possa se diferenciar e comparar uma personalidade de outra. Para que isso seja feito existem diversas abordagens de como se usar e definir a Personalidade, cada uma dessas abordagens foca em diferentes conceitos e medidas, a abordagem que utilizaremos nesse trabalho é a *Trait approach*. Dentro dessa abordagem se destaca, após muitos estudos, que todo humano pode ser classificável/diferenciável por ao menos esses cinco diferentes traços de personalidade chamados de *Big Five*. Se em vez de definir a personalidade unicamente, pudesse dividi-la em várias partes e dizer quanto que cada humano tem de cada uma dessas partes seria mais simples apresentar e comparar as personalidades. E é isso que a *Trait approach* sugere, que a personalidade seja definida em traços que possam ser medidos.

2.1. The Big Five Model

O modelo do Big Five descreve as características humanas de acordo com cinco principais traços: neuroticismo, extroversão, socialização, realização e abertura para o novo, os termos na língua inglesa são: *neuroticism*, *extraversion*, *agreeableness*, *conscientiousness* e *openness to experience*. A combinação destes cinco traços reflete os principais aspectos da personalidade humana. A estrutura do *Big Five* não implica que diferentes personalidades possam ser reduzidas somente a cinco traços. No entanto, essas cinco dimensões são o máximo de abstração da personalidade e cada dimensão resume um grande número de características mais específicas da personalidade [Nunes 2008].

O modelo *Big Five* é um dos modelos mais utilizados propostos até então. O modelo foi proposto não apenas para um entendimento geral do comportamento humano, mas também para profissionais da área de psicologia tratar transtornos de personalidade. Os cinco fatores são considerados a base ou dimensões do espaço da personalidade [Kshirsagar e Thalmann 2007]. As cinco dimensões serão descritas abaixo de acordo com [Hu e Pu 2010].

- Abertura para o novo: apreciação de arte, emoção, aventura, idéias incomuns, curiosidade, e variedade de experiência.
- Realização: tendência a mostrar autodisciplina, agir lealmente, e realizar objetivos. Rígido em vez de um comportamento espontâneo.
- Extroversão: energia, emoções positivas, urgência, e uma tendência a estimular o ambiente na companhia de outras pessoas.
- Socialização: uma tendência para ser compassivo e cooperativo do que desconfiado e antagonista em relação aos outros.
- Neuroticismo: tendência a ter emoções negativas facilmente, como raiva, ansiedade, depressão e vulnerabilidade.

De acordo com [Hu e Pu 2009,2010] o inventário de personalidade que usa cada um desses cinco fatores pode ser ainda subdividido em outros fatores, as quais

chamaram de facetas. No caso dos teste NEO-IPIP descrito posteriormente, cada um dos cinco fatores são divididos em 6 facetas, totalizando 30 facetas. Essas facetas são utilizadas por psicólogos para enriquecer o modelo *Big-Five*, com características mais refinadas [Nunes 2008]. Essas 30 facetas relacionadas com as cinco dimensões serão ilustradas na tabela 1.

Tabela 1. Os cinco grandes fatores e suas respectivas facetas.

Cinco Grandes Fatores	Facetas
Extroversão	Amigabilidade
	Gregarismo
	Assertividade
	Nível de atividade
	Procura por excitação
	Bom humor
Socialização	Confiança
	Moralidade
	Altruísmo
	Cooperação
	Modéstia
	Compaixão
Realização	Auto-eficácia
	Ordem
	Senso de dever
	Empenho
	Autodisciplina
	Prudência
Neuroticismo	Ansiedade
	Raiva
	Depressão
	Autopercepção
	Falta de moderação
	Vulnerabilidade
Abertura	Imaginação
	Interesses artísticos
	Emotividade
	Senso aventureiro
	Intelecto
	Liberalismo

Fonte: [SOUZA e NUNES 2010]

2.2. Extração de personalidade através de inventário -NEO-IPIP

Para extrair os traços de personalidade (como os cinco grandes fatores e suas respectivas facetas) psicólogos geralmente usam inventários. Esses inventários são diretamente aplicados pelos psicólogos, ou eles podem estar disponíveis na *web*. Eles podem ter muitas ou poucas questões, o número de questões no questionário é diretamente relacionado ao quão refinado se deseja extrair os traços de personalidade de

cada pessoa. O questionário que foi usado para validação e teste desse trabalho é o NEO-IPIP [Johnson 2005]. As 300 questões do NEO-IPIP são itens que estão numa escala de cinco pontos. Os pontos são numericamente representados por 1, 2, 3, 4 e 5, associados com a respectiva resposta do usuário. Cada fator do *Big Five Model* é representado por um conjunto de 60 questões, então as 300 questões do NEO-IPIP são os 5 fatores multiplicado por 60 questões de cada um. Essas 60 questões de cada um dos fatores representam 10 questões para cada faceta, então cada fator do Big Five (60 questões) é igual a 6 facetas multiplicado por 10 questões de cada faceta.

3. Extração de personalidade através do ritmo do uso do teclado

Em estratégias baseadas em biometria para identificação e/ou verificação de pessoas, medidas estáticas e/ou dinâmicas são usadas como “senhas” pessoais [Filho e Freire 2006]. Uma dessas medidas é a digitação biométrica, mais comumente referida como ritmo de digitação. Na verdade, o ritmo de digitação analisa a forma como uma pessoa digita as teclas do teclado [Filho e Freire 2006]. Considerando esse fatores, existem duas medidas principais para verificação/identificação baseados no ritmo de digitação, são elas:

- Latência entre teclas consecutivas e
- Tempo que uma mesma tecla permanece pressionada (*hold time*).

A latência deve ser entendida como o tempo entre o apertar de uma tecla e o apertar da tecla seguinte e a segunda medida é o tempo entre o apertar da tecla e o soltar da mesma.

Quando um usuário está digitando três eventos são acionados:

- *KeyDown*: que é acionado imediatamente assim que uma tecla é pressionada.
- *KeyPress*: que é acionado quando uma tecla é pressionada totalmente
- *KeyUp*: que é acionada quando uma tecla é liberada.

Então a captura das medidas de ritmo de digitação é feita da seguinte maneira: quando uma tecla é pressionada o evento *keypress* aciona um timer que é interrompido quando o evento *keypress* é acionado novamente, ou seja, o timer conta o tempo decorrido entre uma tecla pressionada e outra. Esse é o tempo de latência. Já o tempo de tecla pressionada usa os eventos *keydown* e *keyup*, o timer nesse caso é disparado pelo evento *keydown* e interrompido pelo evento *keyup*, ou seja, o timer conta o tempo decorrido entre o apertar e soltar de uma mesma tecla.

Em [Khan 2008] ele analisa se comportamentos podem ser observados como comportamentos interativos, expressa através de um teclado. Obtendo os dados de interação estes podem ser analisados para determinar se existe qualquer ligação entre os traços de personalidade e o comportamento via teclado. Isso proporciona a base para um instrumento para medir a personalidade com base em comportamentos interativos facilitando a medição e determinação do perfil de personalidade do usuário, sem a necessidade do envolvimento explícito do participante, por exemplo, através de inventários.

Em [Porto et al 2011] a extração de traços de personalidade é realizada através do teclado. Obtêm-se informações de como o usuário digita determinado texto, utilizando latência entre teclas consecutivas (o tempo entre o apertar de uma tecla e da subsequente). Tendo a informação de como um usuário digita, este pode ser comparado com um banco de dados que contem informações de outros usuários, informações estas

que são o ritmo de digitação e a personalidade do usuário, e através de técnicas de agrupamento de dados (clustering), o usuário é isolado em um grupo que tem o mesmo padrão de digitação, e então os traços de sua personalidade são inferidas a partir dos usuários que fazem parte do mesmo grupo. Os testes realizados para essa abordagem tendem a confirmar que quando os usuários são agrupados levando em consideração o ritmo de digitação, alguns traços de personalidade são correlacionados. Usuários que são agrupados num mesmo cluster tendem a ter o mesmo valor para algumas facetas, e, portanto usuários que não se soubessem o valor das facetas, tendo sido agrupados em um cluster, poderiam ter seus valores inferidos pelos outros elementos do mesmo cluster.

4. Sistemas de Recomendação

Sistemas de Recomendação começaram a aparecer na década de 90, eles são aplicações que fornecem recomendações personalizadas aos usuários, recomendações de produtos, serviços ou pessoas no quais eles possivelmente estarão interessados. As pessoas comumente tendem a seguir os conselhos dados por outras pessoas. Inicialmente, os Sistemas de Recomendação usavam as informações obtidas através das opiniões de pessoas que já tinham utilizado aquele produto ou serviço, ou se relacionado com a pessoa a ser recomendada [Nunes 2008].

Existem muitas técnicas de recomendação, podemos citar cinco delas, *Content-Based*, *Collaborative Filtering*, *Demographic*, *Knowledge-based*, e *Utility-based*, cada uma com seu ponto a favor e com seu ponto contra. Na técnica do *Content-based*, cada vez que um usuário demonstra interesse em algum item, essa informação é guardada, e no futuro, itens que são similares aos o que o usuário demonstrou interesse no passado, são recomendados ao usuário. Na técnica *Collaborative Filtering*, recomenda-se itens que pessoas com gostos e preferências similares gostaram no passado.

Na *Demographic*, recomendações são feitas baseadas nos dados demográficos do usuário, como idade e gênero. Na *Knowledge-based*, recomendações são feitas a partir de conclusões tiradas a partir das preferências e necessidades do usuário. E, finalmente, na *Utility-based*, recomendam-se os itens baseados na utilidade que estes terão aos usuários. Geralmente sistemas de recomendações utilizam uma técnica híbrida de recomendação. Utilizando uma ou mais das técnicas descritas acima. Normalmente, a técnica *Collaborative Filtering* quando combinada com mais alguma técnica obtém os melhores resultados [Nunes 2008].

Nesse trabalho incita-se o uso de uma técnica diferente das descritas acima, a técnica é a *Personality-based*, que consiste em fazer recomendações baseadas na personalidade do usuário. Como dito na seção 1 desse artigo, Sistemas de Recomendação baseados em personalidade tem sido bem aceitos por pesquisadores educadores e profissionais da indústria, por mostrar resultados interessantes, e porque há comprovações que relacionam personalidade com comportamento. No entanto, como descrito na seção 2.2, a extração da personalidade geralmente é feita a partir de questionários, questionários estes que na maioria das vezes são enormes, como é o caso do NEO-IPIP que contem 300 questões. Dessa forma o ideal é propor outra forma de extração de personalidade menos intrusiva e menos cansativa, empiricamente foi proposto o estudo do uso de informações advindas do ritmo do teclado, como descrito na seção anterior, objetivando buscar alguma correlação com traços de personalidade e

posteriormente usar o teclado como forma de captar informação do usuário/aluno para então personalizar ambientes a ele considerando os aspectos de sua personalidade para tomada de decisão computacional principalmente em ambientes educacionais.

5. Conclusões parciais e discussões

Como destacamos na seção 1, Sistemas de Recomendação baseados em personalidade têm se mostrado eficientes no que diz respeito à satisfação dos usuários com as recomendações a eles dirigidas. Mas como também foi destacada na seção 1 a tradicional forma de se extrair a personalidade de um usuário é através de inventários que na maioria dos casos ultrapassam as 100 questões, como é o caso do inventário NEO-IPIP descrito na seção 2.2 que contém 300 questões.

Na seção 3 foi descrito que existem estudos que sugerem que traços da personalidade podem ser identificados através do uso de diversas ferramentas além do tradicional inventário. Uma dessas ferramentas foi o teclado, onde através do ritmo de digitação aspectos de personalidade são descobertos. O resultado do experimento preliminar com o uso do teclado como descrito detalhadamente em [Porto et al 2011] revela quais são os traços que sugerem uma correlação entre seus valores e o ritmo de digitação. O experimento foi realizado com setenta e duas pessoas. Essa preencheram o inventário NEO-IPIP na sua completude (300 questões) e tiveram seus ritmos de digitação capturados para a formação do banco de dados necessário para o fechamento das correlações. Esse experimento objetivou mostrar quais os traços que estão relacionados com a digitação, então utilizando esse banco de dados de 72 pessoas foram formados grupos considerando seu ritmo de digitação, utilizando a técnica de agrupamento de dados (*clustering*). Após formados os grupos, o próximo passo foi analisar o que a personalidade dos usuários de um mesmo grupo tinham em comum, quais os traços que eram correlacionados com o ritmo de digitação, e finalmente, depois de testes e uma análise estatística chegou-se a conclusão de que duas facetas e um fator poderiam ser extraídos utilizando o ritmo de digitação, são eles: o fator, Extroversão e as facetas Autodisciplina e Interesses Artísticos. (Essa pesquisa faz parte das conclusões preliminares do TCC de curso de graduação dos autores). Testes adicionais com uma base de dados mais densa estão sendo monitorados e podem revelar outras correlações entre facetas e/ou fatores.

A seção 4 finalizou o artigo apresentando de que forma pode-se utilizar os dados extraídos de personalidade, note que uma área potencial de aplicação de Sistemas de Recomendação é a área Educacional.

5.1. Questões para discussão

Esse *position paper* incita sobre a possibilidade de uso de aspectos psicológicos na tomada de decisão computacional. Existem trabalhos que provam que aspectos afetivos, bem como inteligência emocional [Gonzalez 2007] e personalidade tem trazido indícios de adequada personalização na Educação, então o que discute-se é:

- Alunos tendem a se reunir em grupo, dessa forma a personalização na recomendação de grupo pode incentivar ou motivar os alunos à construção colaborativa do conhecimento em Educação convencional? Ou essa personalização seria mais adequada à Educação à distancia devido ao fato de características de presença social estarem ausentes em decorrência da mídia

utilizada? A necessidade de encontrar compatibilidade entre tutores e alunos torna-se tarefa árdua na Educação a Distância? A formação de grupos nesse contexto beneficiaria aspectos da Educação?

- A extração de traços de personalidade objetivando alimentar o modelo do Aluno ocorre efetivamente quando utilizamos um teste de 300 questões? Ou nesse contexto o aluno fica desmotivado pela extensão do inventário, preenchendo erroneamente o teste e dessa forma produzindo um modelo de aluno inconsistente? Seria interessante prover outro tipo de interface que permita extrair as informações de personalidade dos alunos de uma forma menos árdua?
- Os traços de personalidade podem efetivamente auxiliar na personalização do ambiente de ensino aprendizagem? de que forma?

Referências Bibliográficas

- Brinkman, W.-P., Fine, N. (2008) "Personalising emotional skin designs". Emotion in HCI: Joint Proceedings of the 2005, 2006 and 2007 International Workshops, p.31-33, Fraunhofer IRB Verlag Stuttgart.
- Filho, J; Freire, O. (2006) "On the Equalization of Keystroke Time Histograms". Pattern Recognition Letters. Elsevier, Vol. 27, Issue 12, pp. 1440-1446.
- Gonzalez, G., De La Rosa, J.L., And Montaner, M. (2007) Embedding Emotional Context Inrecommender Systems. In The 20th International Florida Artificial Intelligence Research Society Conference-Flairs, Key West, Florida.
- Hu, R., Pu, P. (2009) "Acceptance Issues of Personality-based Recommender Systems". Proceedings of the third ACM conference on Recommender Systems, ACM, New York, NY.
- Hu, R., Pu, P. (2010) "Using Personality Information in Collaborative Filtering for New Users". Recommender Systems and the Social Web.
- Johnson, John A. (2005) "Ascertaining the validity of individual protocols from webbased personality inventories". Journal of research in Personality, 39(1):103-129.
- Khan, I., et al. (2008) "Measuring Personality from Keyboard and Mouse Use". ACM International Conference Proceeding Series, Vol. 369. Portugal.
- Kshirsagar, S., Thalmann, N.M. (2007) "A Multilayer Personality Mode". In Proceedings of the 2nd International Symposium on Smart Graphics, p. 107-115.
- Nunes, M. A. S. N. (2008) "Recommender Systems based on Personality Traits", Tese de Doutorado, Université Montpellier 2, França.
- Nunes, M. A. S. N. ; Cazella, S. C. (2011) "O que sua Personalidade revela? Fidelizando clientes web através de Sistemas de Recomendação e Traços de Personalidade". In: Patricia Vilain e Valter Roesler. (Org.). Tópicos em Banco de Dados e Multimídia e Web. Porto Alegre: SBC, 2011, v. 1, p. 91-122.
- Nunes, M. A. S. N. (2011) "Computação Afetiva personalizando interfaces, interações e recomendações de produtos, serviços e pessoas em ambientes computacionais". In. Portfólio DCOMP. Editora UFS.
- Porto, S. M.; Costa, S. W.; Nunes, M. A. S. N; Matos, L. N. (2011) "Desenvolvimento de Metodologias de Extração de Perfil Psicológico de Usuário para Aplicação em Sistemas de Recomendação Objetivando Personalização de Produtos e Serviços em E-Commerce". Relatório Técnico de Pesquisa. Universidade Federal de Sergipe
- Self, John. (1999) "The Defining Characteristics of Intelligent Tutoring Systems Research: ITS care, precisely". IN: International Journal of Artificial Intelligence in Education, 10, 350-364.
- Souza, D. A. D.; Nunes, M. A. S. N. (2010) "Aspectos teóricos e mensuração do construto psicológico "Personalidade" - Desenvolvimento do plugin Moodle para a formação de grupos de trabalho para uso na EAD UFS". Universidade Federal de Sergipe. São Cristóvão, p. 53.