

Análise de Interações Sociais como Regras se-então em Ambientes Sociais Multimídia

Alan Keller Gomes

Orientadora: Prof^a. Dr^a. Maria da Graça Pimentel
{alankeller,mgp}@icmc.usp.br

RESUMO

Com o avanço das tecnologias Web e dos recursos de computação inseridos em cenários como TV digital e dispositivos móveis, é cada vez maior a necessidade de se projetar e desenvolver artefatos de software capazes de dar suporte a interação social entre os usuários. A análise de interações sociais em Ambientes Sociais Multimídia (ASM) trata da identificação de fatores relacionados e também do estabelecimento de heurísticas e guidelines para a avaliação da sociabilidade. Entretanto, ainda não está estabelecido um modelo para representação e avaliação das interações sociais em ambientes multimídia. Neste trabalho é descrito o foco da minha tese de doutorado, que trata da especificação de um processo de análise de interações sociais em ASM, a partir da representação e avaliação dessas interações sociais como regras *se-então*.

PALAVRAS-CHAVE

Interações Sociais, Ambientes Multimídia, Comportamento do Usuário, Medição

1. INTRODUÇÃO

Este trabalho descreve o foco da minha tese de doutorado que está em desenvolvimento no Programa de Doutorado em Ciência de Computação do Instituto de Ciências Matemáticas e de Computação (ICMC) na Universidade de São Paulo (USP – São Carlos). O trabalho teve início em agosto de 2008 e tem expectativa de término em julho de 2012.

Ambientes Sociais Multimídia permitem que seus usuários possam manipular, buscar e compartilhar objetos multimídia (OM) tais como animações, fotos, músicas, vídeos, etc. O principal objetivo dos ASM é permitir que seus usuários estabeleçam relações sociais baseadas em OM [20]. Os estudos e pesquisas relacionados a interação social em ASM são tratados principalmente no contexto do Software Social [3] e da TV e Vídeo Sociais [4].

As redes sociais são ASM que se destacam pela sua popularidade, e ainda, pelas facilidades de acesso e de interação. Na análise de redes sociais [25], os usuários e suas relações são representados como um grafo [16]. Na análise das relações em redes sociais, geralmente são empregadas técnicas de mineração de dados tais como agrupamento, predição e classificação [15]. Nesse cenário, há uma oportunidade de pesquisa para o provimento de um modelo baseado em regras, que permita a representação e a avaliação de situações envolvendo os usuários de ASM em interações sociais, deixando explícitas as ações executadas pelos usuários.

Resultados da Psicologia Social argumentam que as interações sociais podem ser especificadas como *contingências comportamentais* na forma de regras *se-então*. Como exemplo, ao observar uma interação social, envolvendo os usuários a e b que executam ações A_1 e A_2 tendo como consequência C_1 , essa observação pode ser registrada como a regra $aA_1 \cap bA_2 \rightarrow abC_1$ [19]. No campo de pesquisa do Aprendizado de Regras [15], regras *se-então* são representadas genericamente como implicações, na forma $B \rightarrow H$, que podem ser avaliadas utilizando-se de procedimentos de mineração de dados já estabelecidos e de uma variedade de medidas simétricas e assimétricas de regras [18].

O trabalho aqui apresentado trata da especificação de um processo de análise de interações sociais em ASM, a partir da representação e avaliação dessas interações sociais como regras *se-então*. Na representação é utilizada a linguagem de Mechner para codificação de contingências comportamentais como implicações lógicas. Na avaliação são utilizados procedimentos de mineração de dados e medidas simétricas e assimétricas de regras.

Este trabalho está organizado da seguinte forma: na Seção 2 são apresentados os fundamentos teóricos e os trabalhos relacionados; na Seção 3 são reportadas a justificativa, a questão de pesquisa e os objetivos da tese; na Seção 4 é apresentada a técnica de representação e avaliação de interações sociais proposta na tese; na Seção 5 são apresentadas a metodologia empregada e são reportadas as contribuições obtidas e esperadas; e na Seção 6, as considerações finais.

2. REFERENCIAL TEÓRICO E TRABALHOS RELACIONADOS

O estudo dos aspectos relacionados a sociabilidade em ASM tem o propósito de dar suporte a interação social e aumentar a conectividade social entre os usuários. Esses aspectos são

WebMedia'11: Proceedings of the 17th Brazilian Symposium on Multimedia and the Web. XI Workshop on Ongoing Thesis and Dissertations.

October 3 -6, 2011, Florianópolis, SC, Brazil.

ISSN 2175-9650.

SBC - Brazilian Computer Society

estudados no contexto do Software Social [3] e da TV e Vídeo Sociais [4]. Entretanto, ainda não está estabelecido um modelo para representação e avaliação das interações sociais propriamente ditas em ASM.

Heurísticas e guidelines que dão suporte a avaliação da sociabilidade no contexto de TV e Vídeo Social foram propostas por Geerts e Grooff [9]. Na concepção de aplicações para TV móvel, Geerts [8] e Chorianopoulos [5] investigaram questões relacionadas a sociabilidade. Fatores relacionados a sociabilidade em ASM também são tratados no contexto do Software Social, como apresentaram Gao et al. [7].

Dentre os softwares sociais, destacam-se na atualidade as redes sociais. Na análise de redes sociais [23], os usuários e suas relações são representados como um grafo [16]. Por exemplo, relações simétricas são representadas como um grafo não-direcionado e relações assimétricas são representadas como um grafo direcionado [21]. Nessa análise, geralmente, são empregadas técnicas de mineração de dados para a identificação de grupos (clustering), a predição e a classificação das atividades dos usuários [25].

A representação de relações em redes sociais como um grafo permite o estudo da interação entre os usuários, como proposto por Wilson et al. [24], para dar sentido às interações sociais de usuários do Facebook. Combinando a posição do usuário, polaridade de opinião e de qualidade textual dos tweets, Bigonha e Cardoso [2] propuseram uma técnica para classificação de usuários do Twitter¹.

Técnicas de mineração de dados são usados na análise de redes sociais, por exemplo, em tarefas de identificação de grupos de usuários do Twitter [1]. Modelos de regressão são empregados em tarefas de predição [11], e também em tarefas de classificação [17] de atividades dos usuários. Nesse contexto, identifica-se a necessidade de um modelo baseado em regras que permita a representação e a avaliação das interações sociais em ASM.

Em trabalhos anteriores, estudamos interações sociais associadas ao compartilhamento assíncrono de links de vídeo e de sessões de anotações [12] e também ao compartilhamento síncrono e assíncrono de anotações colaborativas [13] sobre vídeos do YouTube, explorando uma abordagem social do paradigma *Watch-and-Comment* [6]. Agora, aplicamos nossa técnica de representação e avaliação de interações sociais baseadas em mídias entre usuários do Facebook [14].

3. JUSTIFICATIVA E OBJETIVOS

Os modelos existentes para representação e avaliação das interações sociais não são capazes de explicitar quais ações do usuário são executadas e, com essa limitação, fornecem poucos subsídios para que seres humanos interpretem como as interações sociais ocorrem. Nesse cenário, existe a oportunidade de se prover um modelo descritivo para representação e avaliação de interações sociais.

Partindo dessa observação, formula-se a questão de pesquisa que é investigada na minha tese de doutorado: “É possível especificar um processo de análise que permita a represen-

¹www.twitter.com.

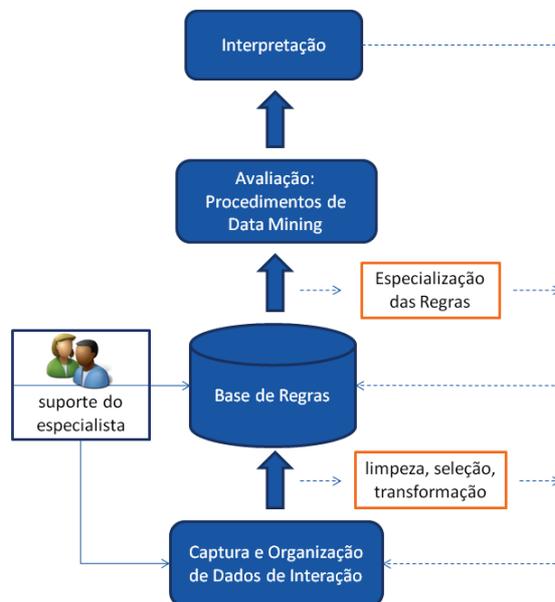


Figura 1: Análise de Interações Sociais em ASM

tação e avaliação de interações sociais em ASM, deixando explícitas as ações do usuário e que possa ser interpretado por seres humanos com alguma facilidade?”.

A partir da formulação desta questão de pesquisa, o objetivo geral da minha tese de doutorado é especificar um processo interativo e iterativo de análise de interações sociais em ASM, a partir da representação e avaliação das interações sociais como regras *se-então*, deixando explícitas as ações executadas pelos usuários. Na Figura 1 é apresentada uma visão geral do processo que almeja-se especificar.

Como objetivos específicos são estabelecidos:

- O1 organizar os dados de interação entre usuários de ASM para que seja possível a sua utilização na avaliação de regras *se-então* codificadas na linguagem de Mechner;
- O2 desenvolver uma técnica para a representação e a avaliação de interações sociais como regras *se-então*, tornando explícitas as ações dos usuários em ASM;
- O3 aprimorar a técnica desenvolvida no objetivo O2, de modo que seja possível a especialização de um conjunto de regras, tornando explicitos também os tipos de OM utilizados, os servidores e/ou aplicações usados para carregar essas mídias, etc;
- O4 desenvolver uma ontologia de interação social para ASM baseada em regras *se-então*.

A metodologia empregada na elaboração da tese para que o objetivo geral e os específicos sejam atingidos, é apresentada na Seção 5.

4. TÉCNICA PROPOSTA NA TESE

Nesta seção é apresentada a técnica desenvolvida para a representação e a avaliação de interações sociais em ASM.

4.1 A Linguagem de Mechner

Em ciências sociais, as interações diárias entre as pessoas, ou seja, quaisquer tipos de interações sociais, podem ser especificadas como *contingências comportamentais* [22]. Mechner [19] apresentou uma linguagem formal simbólica para codificar quaisquer contingências comportamentais como implicações lógicas. Alguns elementos importantes dessa linguagem são:

1. *Ação (ou ações)*: combinando o antecedente da contingência, ou seja, $A \rightarrow$. Havendo mais de uma ação, elas são representadas como $A_1 \cap A_2 \cdots \rightarrow$.
2. *agente(s) da ação(ões)*: representados por letras minúsculas e colocado na frente de um A . Por exemplo, agente a que realiza a ação A é representado como aA . Uma letra pode representar um único agente ou um grupo de agentes que realizam a ação.
3. *Consequências*: Que corresponde ao consequente da contingência, ou seja, $\rightarrow C$. Havendo mais de uma consequência, são representadas como $\cdots \rightarrow C_1 \cap C_2$.

Por exemplo, uma contingência comportamental codificada na linguagem de Mechner é: $aA_1 \cap bA_2 \rightarrow abC_1 \cap abC_2$, ou seja, *se a executar a ação A_1 e b executar a ação A_2 então as sequencias C_1 e C_2 são percebidas por a e b .*

Embora outros sistemas de notação já tenham sido propostos, optou-se pela linguagem de Mechner [19] porque esta permite a descrição de contingências comportamentais como implicações lógicas na forma normal disjuntiva, que é exigida pelos procedimentos de mineração de dados adotados.

4.2 Representação e Avaliação de Interações

Na minha tese proponho o uso da linguagem de Mechner para representar situações envolvendo usuários em interações sociais nos ASM, identificando-se *ações*, *agentes de ações* e *consequências*. Dessa forma, essas situações podem ser representadas como, por exemplo,

- *se* um usuário do Facebook a executa a ação $A_1 =$ *postar uma mensagem de texto* em seu mural,
 - *então* usuário a e usuários nos grupos k e l $C_1 =$ *são notificados desta postagem*,
 - *se* os usuários no grupo k realizam a ação $A_2 =$ *comentário de um post* (depois de notificado da ação do usuário a),
 - * *então* o usuário a e usuários em grupos k e l $C_2 =$ *são notificados deste comentário*,
- *então* usando a linguagem de Mechner, esta interação social é representada como $aA_1 \cap kA_2 \rightarrow akC_1 \cap akC_2$.

A ação que começa a interação social, como é o caso de A_1 no exemplo acima, é chamada de *estímulo social* [22]. De modo análogo, interações sociais baseadas em mais de um tipo de mídia, como no Twitter e Google+², e baseadas em um tipo de mídia, como no YouTube e Soundcloud, podem ser representadas.

²www.googleplus.com

Na minha tese proponho que contingências comportamentais sejam genericamente representadas como implicações $Corpo \rightarrow Cabeça$ ($Body \rightarrow Head$), em suma, $B \rightarrow H$ (regra R). Por exemplo, considerando $R = aA_1 \cap kA_2 \rightarrow akC_1 \cap akC_2$ como uma contingência comportamental, $B = aA_1 \cap kA_2$ e $H = akC_1 \cap akC_2$.

O objetivo é que um conjunto de comportamentos (observados em um ambiente social e formalizados como contingências na linguagem de Mechner) seja mapeado em um conjunto de regras típicas da mineração de dados, e em seguida, procedimentos já estabelecidos de mineração de dados sejam usados para avaliar e medir o conjunto de comportamentos.

Usando técnicas de mineração de dados, uma implicação $B \rightarrow H$ pode ser avaliada comparando-a com um conjunto de observações [18]. Por exemplo, o número de n de contingências comportamentais observadas durante uma determinada experiência social pode ser calculado usando os valores clássicos da avaliação de regras de modo que $n = bh + b\bar{h} + \bar{b}h + \bar{b}\bar{h}$.

Esses valores são: bh , número de situações observadas nas quais o corpo b e a cabeça h são verdadeiros; $b\bar{h}$, número de situações observadas nas quais o corpo b é verdade e a cabeça h é falsa; $\bar{b}h$, número de situações observadas nas quais o corpo b é falso e a cabeça h é verdade; $\bar{b}\bar{h}$, número de situações observadas nas quais o corpo b e a cabeça h são falsos.

Com o mapeamento de contingências na linguagem de Mechner para regras de mineração de dados, valores típicos da avaliação de regras podem ser usados para calcular medidas de avaliação [10], como as apresentadas na Tabela 1.

Tabela 1: Medidas de Avaliação de Regras

Suporte	Correlação	Confidência	Sensitividade
$\frac{bh}{n}$	$\frac{bh}{n * \sqrt{\frac{b * h}{n^2}}}$	$\frac{bh}{bh + b\bar{h}}$	$\frac{bh}{bh + \bar{b}h}$

Medidas de avaliação regra podem ser *simétricas* ou *assimétricas*. As medidas de suporte e correlação são simétricas, pois seus valores são idênticos para $B \rightarrow H$ e $H \rightarrow B$. No entanto, as medidas de confidência e sensitividade são assimétricas, pois seus valores podem não ser os mesmos para as regras $B \rightarrow H$ e $H \rightarrow B$. Por convenção, os valores destas medidas são entre 0% e 100% em vez de entre 0 e 1 [15].

5. METODOLOGIA E CONTRIBUIÇÕES

A metodologia empregada na elaboração da minha tese de doutorado apresenta os seguintes passos:

1. **Revisão de Literatura:** Nessa atividade são investigados trabalhos tratam da análise de relações e interações sociais em ASM e da utilização de técnicas de mineração de dados nessa análise;
2. **Coleta e Organização de Dados de Interação:** A coleta de dados é feita de forma automática, por exemplo, utilizando-se API's ou analisando-se log's gerados a partir da utilização de ASM. Nessa fase, o suporte de um especialista é necessário para que possam ser

identificados, capturados e armazenados os dados relacionados aos elementos da linguagem de Mechner;

3. **Representação e Avaliação de Interações Sociais:** É feita de acordo com a técnica apresentada na sessão anterior. Quando uma nova regra é codificada, o suporte do especialista é necessário para validar a regra;
4. **Interpretação dos Resultados e Aprimoramentos:** Os resultados obtidos a partir da avaliação e mensuração das regras são submetidos a interpretação de usuários interessados na análise das interações sociais. A partir dessa interpretação, ajustes e aprimoramentos podem ser feitos nas etapas anteriores.

Foram reportados na literatura alguns resultados preliminares a partir da análise de interações sociais em um ASM baseado em Vídeo Social [13] [12]. Atualmente, a validação da especificação do processo está sendo feita a partir da extração de dados e análise de interações sociais do Facebook.

A principal contribuição esperada da minha tese de doutorado é a especificação de um processo de análise de interações sociais em ASM, que explicita as ações executadas pelos usuários, a partir da representação e avaliação das interações sociais como regras *se-então*. Obtendo-se a contribuição principal ao se atingir o objetivo geral, espera-se que os objetivos específicos também tenham sido atingidos, e assim, as demais contribuições:

- C1 uma organização dos dados de interação social em ASM, para a coleta e mineração desses dados;
- C2 uma técnica de representação, avaliação e a especialização de interações sociais em ASM representadas como regras *se-então*;
- C3 uma ontologia de interação social em ASM baseada em regras *se-então*.

No estado atual da tese, a contribuição C2 foi reportada na literatura [14]. As contribuições C1 e C3 estão sendo apreciadas pela comunidade científica. Finalizados alguns ajustes e aprimoramentos, que também serão reportados, espera-se que a contribuição principal seja obtida.

6. CONSIDERAÇÕES FINAIS

Neste trabalho apresento o foco de pesquisa da minha tese de doutorado que é a especificação de um processo de análise de interações sociais em ASM, a partir da representação e avaliação dessas interações sociais como regras *se-então*. A representação de interações sociais como um conjunto de contingências comportamentais (regras *se-então*) é capaz de explicitar as ações realizadas pelos usuários. Na avaliação de interações sociais são utilizados procedimentos de mineração de dados na computação de medidas simétricas e assimétricas de regras, que por sua vez são empregadas na avaliação de relações simétricas e assimétricas entre usuários de ASM. A especificação de um processo de análise de interações sociais pode ser muito útil tanto para analistas, desenvolvedores e avaliadores de ASM, como também aqueles interessados na análise das interações sociais que ocorrem nesses ambientes.

Agradecimentos: CAPES, CNPq, FAPESP, FINEP e MCT.

7. REFERÊNCIAS BIBLIOGRÁFICAS

- [1] S. Abrol and L. Khan. Tweethood: Agglomerative clustering on fuzzy k-closest friends with variable depth for location mining. In *IEEE Inter. Conf. on Social Computing (SocialCom'10)*, pages 153–160, 2010.
- [2] C. A. S. Bigonha, T. N. C. Cardoso, M. M. Moro, V. A. F. Almeida, and M. A. Gonçalves. Detecting evangelists and detractors on twitter. In *ACM Braz. Symp. on Web and Mult. (WebMedia)*, pages 107–114, 2010.
- [3] D. M. Boyd and N. B. Ellison. Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13(1):11, 2007.
- [4] K. Chorianopoulos. User interface design principles for interactive television applications. *Int. J. Hum. Comput. Interaction*, 24(6):556–573, 2008.
- [5] K. Chorianopoulos. Scenarios of use for sociable mobile tv. In A. Marcus, R. Sala, and A. C. Roibs, editors, *Mobile TV: Customizing Content and Experience*, Human-Computer Int. Series, pages 243–254. Springer London, 2010.
- [6] R. Fagá, Jr., V. G. Motti, R. G. Cattelan, C. A. C. Teixeira, and M. G. C. Pimentel. A social approach to authoring media annotations. In *ACM Symp. on Doc. Engineering (DocEng'10)*, pages 17–26, 2010.
- [7] Q. Gao, Y. Dai, Z. Fan, and R. Kang. Understanding factors affecting perceived sociability of social software. *Computers in Human Behavior*, 26(6):1846–1861, 2010.
- [8] D. Geerts. The sociability of mobile tv. In A. Marcus, R. Sala, and A. C. Roibs, editors, *Mobile TV: Customizing Content and Experience*, Human-Computer Int. Series, pages 25–28. Springer London, 2010.
- [9] D. Geerts and D. D. Grooff. Supporting the social uses of television: sociability heuristics for social tv. In *ACM Int. Conf. on Human Factors in Comp. Sys. (CHI)*, pages 595–604, 2009.
- [10] L. Geng and H. J. Hamilton. Interestingness measures for data mining: A survey. *ACM Comp. Survey*, 38:1–32, 2006.
- [11] S. Golder and S. Yardi. Structural predictors of tie formation in twitter. In *IEEE Int. Conf. on Social Computing (SocialCom'10)*, pages 88–95, 2010.
- [12] A. K. Gomes, D. C. Pedrosa, and M. G. C. Pimentel. Evaluating asynchronous sharing of links and annotation sessions as social interactions on internet videos. In *IEEE Int. Symp. Applications and the Internet (SAINT)*, 2011.
- [13] A. K. Gomes and M. G. C. Pimentel. Measuring synchronous and asynchronous sharing of collaborative annotations sessions on ubi-videos as social interactions. In *IEEE Int. Conf. on Ubi-media Comp. (U-Media)*, 2011.
- [14] A. K. Gomes and M. G. C. Pimentel. A technique for human-readable representation and evaluation of media-based social interactions in social networks. In *ACM Braz. Symp. on Web and Mult. (WebMedia)*, TO APPEAR, 2011.
- [15] J. Han and M. Kamber. *Data Mining: Concepts and Techniques, 2nd Ed.* Morgan Kaufmann, 2005.
- [16] P. W. Holland and S. Leinhardt. A method for detecting structure in sociometric data. *American Journal of Sociology*, 76(3):pp. 492–513, 1970.
- [17] D. Irani, S. Webb, C. Pu, and K. Li. Study of trend-stuffing on twitter through text classification. In *Micro. Conf. on Coll., Elec. Mess., Anti-Ab. and Spam (CEAS'10)*, 2010.
- [18] N. Lavrac, P. A. Flach, and B. Zupan. Rule evaluation measures: A unifying view. In *Int. Work. on Ind. Logic Prog. (ILP'99)*, pages 174–185. Springer, 1999.
- [19] F. Mechner. Behavioral contingency analysis. *Behavioral Processes*, 78(2):124–144, 2008.
- [20] K. Musial, P. Kazienko, and T. Kajdanowicz. Multirelational social networks in multimedia sharing systems. In N. e. a. Nguyen, editor, *Knowl. Proc. and Reas. for Infor. Society*, pages 275–292. Academic Pub. House EXIT, 2008.
- [21] D. M. Romero and J. M. Kleinberg. The directed closure process in hybrid social-information networks with an analysis of link formation on twitter. In *Int. Conf. on Weblogs and Soc. Media (ICWSM'10)*, 2010.
- [22] B. F. Skinner. *Science and Human Behavior*. N.Y. Press, 1953.
- [23] S. Wasserman and K. Faust. *Social Network Analysis: Methods and Applications*. Cambridge Uni. Press, 1994.
- [24] C. Wilson, B. Boe, A. Sala, K. P. Puttaswamy, and B. Y. Zhao. User interactions in social networks and their implications. In *ACM European Conf. on Computer Systems (EuroSys '09)*, pages 205–218, 2009.
- [25] G. Xu, Y. Zhang, and L. Li. *Web Mining and Social Networking: Techn. and Appl.* Springer-Verlag, 2010.