

Estratégia de orientação para sistemas de hipermídia por meio de navegação colaborativa

Diogo Cezar Teixeira Batista
xgordo@gmail.com

Ligia Flávia Antunes Batista
ligia@utfpr.edu.br

Universidade Tecnológica Federal do Paraná - UTFPR
Av. Alberto Carazzai, 1640
Cornélio Procópio, PR, Brasil

ABSTRACT

This paper presents a model of orientation to users of hypermedia systems based on Ant Colony Optimization theory (ACO). This theory was adapted to used as a way of collaborative navigation. The model was implemented in a prototype and some tests were made. The first results indicate that the model is useful to help users to find the information that they search for.

Categories and Subject Descriptors

H. [Information Systems]: Information Interfaces and Presentation; H.5 [Information Interfaces and Presentation]: Hypertext/Hypermedia—*user issues, navigation*

General Terms

Human factors, Experimentation

Keywords

Adaptative Hypermedia, Orientation, Ant Colony Optimization, Collaborative Navigation.

1. INTRODUÇÃO

O número de sistemas hipermídia existentes implica possibilidade de acesso imediato a imensuráveis quantidades de informação, distribuídas desorganizadamente na Internet. Desta forma, justamente pela diversidade e infinidade de documentos disponíveis na *Web*, torna-se muitas vezes difícil para diversos usuários encontrar a informação desejada. As buscas realizadas geralmente trazem, em meio ao assunto procurado, muito conteúdo não relacionado ou irrelevante. Além disto, sistemas de busca funcionam apenas como filtro para o assunto desejado, sem oferecer alguma forma de assistência navegacional que oriente o usuário até seu objetivo em um determinado sítio, gerando problema de desorientação, pela sensação de não saber a direção a ser seguida. A hipermídia adaptativa é uma área que aborda

a modificação do sistema *Web* de forma a atender as necessidades de cada indivíduo, por meio da adaptação, com métodos específicos para minimizar a desorientação.

Neste contexto, visualiza-se a proposta de navegação colaborativa como uma forma de ajuda mútua entre os usuários que estiverem navegando pelas mesmas páginas. Parte-se da premissa de que grande parte dos usuários percorrem os mesmos caminhos durante a navegação. Para modelar tal solução, baseou-se na teoria do comportamento das formigas, uma vez que a colônia trabalha em grupo e esta característica determina o sucesso de suas tarefas, notadamente a busca e transporte de alimento. Esta teoria foi proposta originalmente para resolver problemas específicos de otimização combinatorial [3, 6, 1] e tem sido utilizada para outras classes de soluções, como processamento de imagens, engenharia de software, eletro-eletrônica, medicina, entre outras [4, 8].

Pode-se fazer uma analogia do comportamento dos usuários que navegam em sistemas *Web* com formigas, em que a comunicação se dá por meio de pistas (feromônio) deixadas em caminhos já percorridos [7]. Se um caminho for muitas vezes percorrido, assume-se que é um trajeto interessante, que deve ser transmitido aos demais. Desta forma, haveria uma comunicação indireta entre os membros da colônia a fim de que todos alcancem seus objetivos, o que demonstra que o sucesso da colônia de formigas em encontrar o alimento advém de uma forma cooperativa de trabalho.

Entretanto isto, ainda não ocorre na Internet, pois os usuários não interagem para colaborar entre si na navegação. Com base nisto, o presente trabalho tem como objetivo modelar uma estratégia de orientação aos usuários por meio de navegação colaborativa. Os fundamentos em que se baseia a abordagem proposta estão detalhados na seção 2. O modelo desenvolvido é apresentado na seção 3. Os resultados observados nos testes são discutidos na seção 4. A seção 5 relata as considerações finais deste trabalho.

2. ABORDAGEM DE ORIENTAÇÃO

A capacidade de auto-organização de insetos sociais é denominada estigmergia, termo também aplicado ao comportamento de cardumes, bandos e enxames. Esta característica é apresentada como um novo paradigma de heurísticas construtivas, *Ant Colony Optimization* (ACO) [3], inspirando trabalhos que produzem soluções otimizadas para problemas como o do caixeiro viajante e roteamento de pa-

cotes .

As formigas reais comunicam-se indiretamente por uma substância chamada feromônio, deixado no meio ambiente. Essa substância é exalada espontaneamente por todos os membros da colônia e possibilita que uma formiga escolha entre dois caminhos diferentes por uma probabilidade baseada na quantidade de feromônio presente em cada um. Logo, se há maior concentração de feromônio em determinado caminho, será este o caminho escolhido [4], gerando um comportamento denominado autocatalítico [3].

No início do processo, quando um grupo de formigas está em busca de alimento e ainda não há um caminho traçado, aquelas que optarem pelo trajeto mais curto irão chegar ao alimento antes das outras. Desta forma, vão também retornar primeiro ao ninho e com isto, acentuar com feromônio a rota percorrida. Com tempo, o hormônio acumula-se mais rapidamente na trilha mais curta, tornando-a a rota preferencial dos membros da colônia [4]. O feromônio evapora com o passar do tempo. Assim, um caminho que não é constantemente visitado e reforçado, tende a desaparecer, deixando de ser o selecionado pelas formigas.

Este trabalho tem seu alicerce nas técnicas de hipermídia adaptativa, as quais são brevemente tratadas na seção 2.1. A aplicação da teoria bio-inspirada do comportamento das formigas aplicada em navegação de sistemas *Web* é apresentada na seção 2.2.

2.1 Hipermídia Adaptativa

A hipermídia adaptativa tem como objetivo proporcionar conteúdo adequado ao perfil ou modelo de cada usuário. Para tal, dois componentes básicos são necessários: a máquina de adaptação, que possui as regras a serem executadas de modo a modificar apropriadamente o que será apresentado, e o modelo do usuário, que apresenta as características, interesses e objetivos daquele que for utilizar o sistema em cada momento.

A adaptação pode se dar em dois espaços distintos: apresentação e navegação. O espaço de apresentação compreende a modificação do conteúdo que será apresentado a cada usuário, enquanto que o de navegação altera a máquina abstrata do hipertexto, ou seja, as ligações (*hyperlinks*) existentes nas páginas que compõem o hiperdocumento [5], contexto em que insere-se esta proposta.

Os métodos existentes para navegação adaptativa indicam o que deve ser feito para proporcionar auxílio ao usuário. São definidos como uma generalização das técnicas existentes, sendo baseados em uma idéia clara de adaptação que pode ser apresentada em nível conceitual [2]. São assim enumerados [5]:

- Condução global: o sistema conduz o usuário na direção de seu objetivo, apresentando uma sequência de passos;
- Condução local: sugere-se um passo de navegação que seja interessante, conforme os dados do modelo do usuário;
- Suporte à orientação global: auxilia o entendimento da estrutura de todo o hiperdocumento;
- Suporte à orientação local: ajuda na compreensão de seu posicionamento no hiperespaço, apresentando informações sobre os nós que podem ser acessados a partir do atual.

As técnicas de navegação adaptativa apontam como deve ser feito o auxílio ao usuário. Fazem parte do nível de implementação em um sistema e podem ser caracterizadas por um tipo específico de representação do conhecimento e por um algoritmo de adaptação [2]. Cita-se [5]:

- Orientação direta: indica em cada ponto da navegação qual o melhor nodo (página) a ser visitado a seguir. Como desvantagem, esta técnica não oferece suporte no caso do usuário não desejar seguir a sugestão do sistema;
- Classificação: ordena os links de um nodo de acordo com sua relevância;
- Ocultação: consiste em restringir o espaço de navegação, ocultando os links para nodos não relevantes para cada usuário;
- Anotação: tem como característica básica aumentar a informação existente nos links com alguma forma de anotação ou comentário que pode dizer mais sobre o estado corrente dos nodos a que se conectam;
- Geração de links: consiste em criar links adicionais automaticamente nos nodos. Uma das desvantagens desta técnica é o aumento no número de links gerados nos nodos, aumentando a sobrecarga cognitiva e tornando o hiperespaço mais complexo;
- Mapas adaptativos: têm o objetivo de ajudar o usuário a entender a estrutura do hiperespaço.

2.2 Aplicação da teoria da colônia de formigas

O relacionamento entre as técnicas de hipermídia adaptativas para navegação e a abordagem comportamental dos insetos sociais se dá fazendo analogia do indivíduo que acessa a *Web* com um certo objetivo e a formiga que percorre certas trilhas em busca de alimento. A sequência de páginas acessada é comparada ao caminho percorrido. Neste contexto, a representação do sistema de hipertexto é a de um grafo, em que as páginas são os vértices e os *hyperlinks* são as arestas.

Uma abordagem anteriormente proposta, implementada no sistema *AntWeb* [7], uma matriz bidimensional representa as taxas de feromônio de cada nodo para todos os outros. A quantidade de feromônio depositada em cada célula da matriz corresponde ao quão relevante uma página destino é perante uma página origem (corrente), considerando um objetivo final como página alvo. Ou, de outra forma, é um indicativo de resposta para a pergunta: estando no nodo i , dar o passo para j , é uma boa escolha a fim de alcançar a página destino d ?

Para isto, faz-se o cálculo da probabilidade de cada formiga escolher ir para a página j , estando na página i na iteração p , a partir da matriz de feromônio, utilizando uma heurística de custo baseada em tempo de obtenção e visita da página j .

Uma das limitações da abordagem *AntWeb* é que não é possível fazer testes com usuários reais, apenas com caminhos aleatórios de navegação que simulam o comportamento de formigas. Além disto, há a necessidade de determinar, para cada usuário, uma página objetivo, a partir da qual ele é conduzido, informação esta difícil de ser determinada, uma vez que muda com o tempo e o que é de interesse de uma pessoa hoje, amanhã pode não ser [7].

Na abordagem ora proposta, denominada OriAnt (Orientação por Formigas), tenta-se contornar estes problemas eliminando a necessidade de determinar a página objetivo de cada usuário e introduzindo a navegação colaborativa para comunicação indireta entre usuários da mesma categoria, tornando possível também a realização de testes com usuários reais. Para isto desenvolveu-se um modelo $M < P, G, F >$ constituído de três conjuntos de dados, representando páginas (P), grupos (G) e feromônio (F). Cada elemento possui seus subelementos de tal forma que:

- P é uma quádrupla $P < p, u, e, c >$ com os seguintes dados de todas as páginas: identificador, data e hora do último acesso, endereço (URL - *Uniform Resource Locator*) e número de acessos, respectivamente;
- G é uma dupla $G < g, n >$, sendo g o identificador do grupo e n o número de acessos daquele grupo;
- F é uma quádrupla $F < o, d, g, qf >$, tendo o como identificador de uma página origem, d da página destino, g como identificador de grupo e qf como quantidade de feromônio.

Desta forma, a camada de navegação adaptativa OriAnt mantém informações de todas as páginas de um sítio *Web* em P. Os interesses dos usuários são mapeados por meio de grupos em G, como uma forma de modelagem de objetivos simplificada, uma vez que o objetivo principal deste trabalho é verificar a aplicabilidade da estratégia de orientação e não a modelagem do usuário. A representação de feromônios é feita em F, uma matriz tridimensional, indicando a relevância da página destino d para o grupo g .

A cada acesso, ocorre o acréscimo de feromônio, conforme a Equação 1. A fórmula de subtração (evaporação) de feromônio, por sua vez, é baseada no conceito de juros compostos, sendo definida pela Equação 2.

$$F_{odg} = F_{odg} + \xi \quad (1)$$

onde:

ξ é uma constante definida pelo administrador do sistema, que significa a relevância de um acesso;

F_{odg} é quantidade de feromônio na aresta que liga a origem o ao destino d para o grupo g .

$$F_{odg} = F_{odg} * (1 - \frac{\varphi}{100})^\tau \quad (2)$$

φ é a taxa de evaporação de feromônio, constante definida pelo administrador, dependente do tempo;

τ é o intervalo de tempo que a página ficou sem acessos, determinado pela equação 3.

$$\tau = t_{atual} - t_{ultimoacesso} \quad (3)$$

3. CAMADA DE ADAPTAÇÃO

Para implementar o modelo descrito na seção 2.2, optou-se por projetar uma camada de adaptação que pudesse ser acoplada a qualquer sistema *Web* para proporcionar navegação adaptativa. A solução proposta OriAnt é atualizada concomitantemente a cada clique e oferece novas orientações para auxiliar o usuário.

A arquitetura do modelo proposto está apresentada na Figura 1, em que as setas apresentam a seqüência de ações que ilustra o comportamento adaptativo da camada OriAnt a partir da navegação de um usuário, que compreende:

1. Escolher o grupo: o usuário faz a opção de seu tema de interesse, para que a partir desta, a camada de adaptação atue. Os grupos de interesse com maior número de visitas recentes aparecem destacados na camada com fontes maiores, conforme ilustrado na Figura 2. Este dado é obtido a partir do sub-elemento n de G. Ressalta-se que os grupos são dinâmicos e mantidos pelo administrador do sistema;
2. Atualizar o banco de dados OriAnt com o dado do grupo, especificando aquele que o usuário tem interesse e incrementando o sub-elemento n de G;
3. Navegar acionando o *hyperlink* desejado;
4. Visualizar a página destino: a camada de orientação é transparente para o usuário e portanto, a página destino é visualizada normalmente após um clique;
5. Realizar a computação de feromônio: calcular a nova quantidade de feromônio para o elemento F_{odg} conforme a equação 1;
6. Atualizar o banco de dados com o novo valor de F_{odg} (página acessada);
7. Efetuar a dedução de feromônio (evaporação): calcular a nova quantidade de feromônio para toda a matriz de feromônio conforme a equação 2;
8. Atualizar o banco de dados com os novos valores de F_{odg} para toda a matriz de feromônio;
9. Iniciar a computação dos links que devem ser sugeridos conforme cálculo de relevância da página, definido pela equação ??

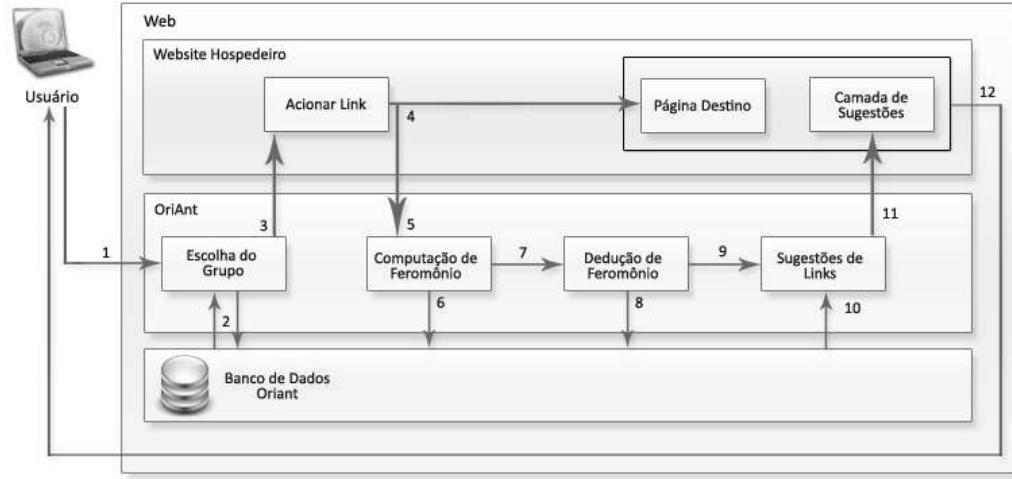


Figure 1: Arquitetura do modelo proposto

10. Obter os dados do banco com relação à navegação dos outros usuários para sugerir links;
11. Exibir sugestões de links conforme a estratégia de orientação selecionada;
12. Aguardar que o usuário acione outro *hyperlink*.

$$\omega(\alpha, \beta) = \frac{F_{\alpha p \gamma}}{\sum_{d=1}^n F_{\alpha d \gamma}} \quad (4)$$

Enfatiza-se que, caso o usuário não faça a escolha do grupo de interesse, o laço de adaptação não é executado. Caso ocorra troca de grupo de interesse durante a navegação, as sugestões passarão a ser computadas com base no novo grupo selecionado.

Além da informação do grupo de interesse, escolhe-se o tipo de orientação (entre objetiva, orientada e relacionada) e o contexto da orientação (esta página ou todas as páginas).

Dependendo do tipo de orientação selecionada a camada OriAnt oferece:

- Disposição objetiva: Mostrará qual é a página alvo. Essa é a página que possui mais feromônio (de um certo grupo). Essa disposição implementa a técnica de orientação direta de [5];
- Disposição orientada: Mostrará qual o caminho que deve-se seguir para chegar até a página alvo. Esse caminho é traçado com base no histórico dos caminhos percorridos de usuários do mesmo grupo; EXEMPLO AQUI
- Disposição por assuntos relacionados: Mostrará quais os links mais visitados pelos usuários de seu grupo. Além da página alvo outras páginas também podem interessar o usuário de determinado grupo, assim ao

selecionar a exibição por assuntos relacionados, obtém-se uma lista das páginas mais relevantes para determinado. Uma ilustração de como a camada se comporta neste tipo de orientação é apresentada na Figura 3.

Dependendo do contexto de orientação, a camada de adaptação comporta-se da seguinte forma:

- Essa Página: Quando esta opção está selecionada a disposição toma como contexto a página atual, ou seja, a partir da página atual, calcula qual é a próxima página mais relevante;
- Todas as Páginas: Quando esta opção está selecionada a disposição toma como contexto todas as páginas do sistema, ou seja, dentre todas as páginas qual é a página mais relevante. Este contexto é particularmente interessante para usuários que não conhecem a estrutura global do site em que estão navegando e já encontram-se desorientados.

A partir do cálculo da relevância de cada página em relação a atual, do tipo de orientação e do contexto, OriAnt sugere páginas, caminhos ou assuntos considerados mais apropriados para o grupo de interesse do usuário.

4. RESULTADOS E DISCUSSÃO

Para realizar testes e avaliar os resultados, foi desenvolvido um site de notícias ao qual foi acoplada a camada de orientação OriAnt. Inicialmente, realizou-se um teste de acompanhamento dos valores da matriz de feromônio, a fim de que fosse possível verificar se a quantidade estava sendo corretamente adicionada em cada clique e a evaporação ocorrendo conforme o desejado. Para isto foi observado se as páginas recentemente acessadas não sofriam alterações de feromônio e se a dedução estava sendo efetuada proporcionalmente ao intervalo de tempo entre os cliques. Os resultados estão apresentados na Tabela 1, em que a coluna nomeada por

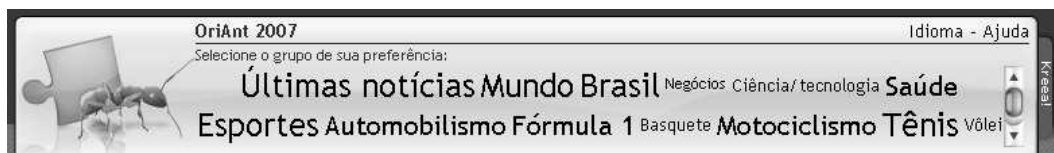


Figure 2: Grupos de interesse

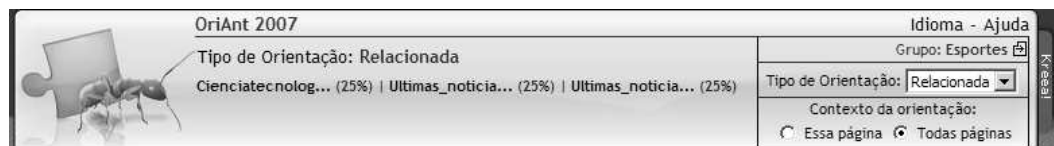


Figure 3: Orientação

Table 1: Evolução da matriz de feromônio

#	origem-destino-grupo	qf0	qf1	qf2	qf3	qf4	qf5
1	1-1-1	10	9.9285	9.8571	9.7856	9.7142	9.6428
2	1-2-1	20	19.8571	19.7143	19.5713	19.4284	19.2856
3	1-3-1	10	9.9297	9.8594	9.7891	9.7187	9.6484
4	1-4-1	10	9.9309	9.8617	9.7925	9.7233	9.6540
5	1-5-2	10	9.9332	9.8663	9.7994	9.7324	9.6653
6	5-6-2	20	19.8664	19.7327	19.5988	19.4648	19.3307
7	5-7-2	10	9.9343	9.8686	9.8028	9.7369	9.6710
8	5-8-2	10	9.9343	9.8686	9.8028	9.7369	9.6710
9	1-9-3	10	20	19.9989	19.9967	19.9935	19.9893
10	9-10-3	20	19.873	19.7466	19.6194	19.4921	19.3646
11	9-11-3	10	9.9378	9.8756	9.8131	9.7506	9.688
12	9-12-3	10	9.9378	9.8756	9.8131	9.7506	9.688
13	9-13-3	0	0	10	9.9994	9.9983	9.9967
14	9-14-3	0	0	0	10	9.9995	9.9985
15	9-15-3	0	0	0	0	10	9.9994
16	9-16-3	0	0	0	0	0	10

origem-destino-grupo representa respectivamente os identificadores da página atual, da página que é o destino do *link* acessado e do grupo de interesse do usuário. A coluna qf0 é a quantidade de feromônio com que cada item da tabela foi inicializado e as demais colunas qf1 a qf5 também indicam a presença de feromônio em cada uma das iterações subsequentes, tendo como iteração um clique do usuário.

Parâmetros de administração aqui (tabelas xls)

As linhas de 13 a 17 representam caminhos que iniciaram sem feromônio, ou seja, nunca haviam sido percorridos. As linhas 1, 3, 4, 5, 7, 8, 9, 11 e 12 apresentam comportamento idêntico, porque a quantidade de feromônio inicial é idêntica e o intervalo de tempo é irrelevante para apresentar uma diferença significativa (número inteiro).

A diferença entre as linhas 2, 6 e 10, que começaram com 20 unidades, deve-se a diferença de tempo entre o acesso das arestas, a linha 2 foi acessada as 12:31:48, a linha 6 as 12:35:35 e a linha 10 as 12:36:29. Pode-se observar que a linha 2 mantém elevada a taxa de feromônio, isso ocorre pois ela foi acionada pelo usuário as 13:11:50.

As linhas de 13 a 17 representam a sequência de páginas acessadas pela primeira vez a partir de um dado destino e

por isto têm seus valores iniciados em 10 em cada iteração.

A Figura 7 apresenta o gráfico da evolução das páginas 1, 2, 10 e 15.

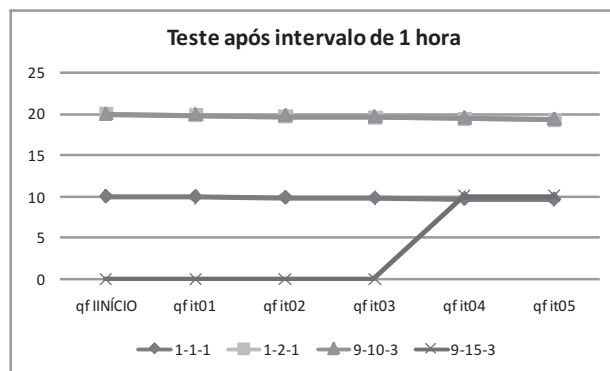


Figure 4: Gráfico dos testes após intervalo de 1 hora

Teste de 10 horas de intervalo: valores iniciais de feromônio, simula-se uma passagem de tempo de 10 horas como se ninguém tivesse acessado o site

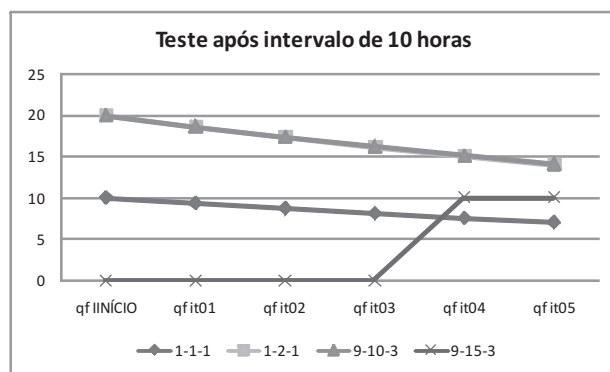


Figure 5: Gráfico dos testes após intervalo de 10 horas

Teste de 24 horas

pra se ter uma 6 casa decimal ou seja x,0000AQUI eh preciso

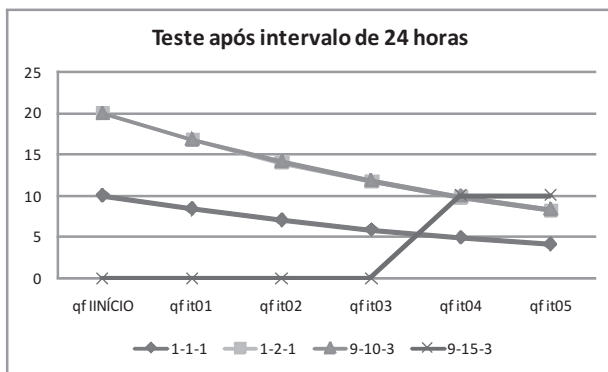


Figure 6: Gráfico dos testes após intervalo de 24 horas

mais que 10 minutos

são as mesmas para os tres casos? 1h, 10 h e 24 h? as diferenças sim...

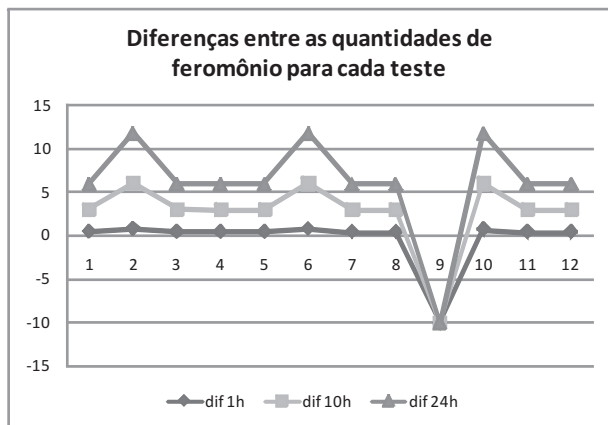


Figure 7: Diferenças ...

5. CONSIDERAÇÕES FINAIS

É possível uma colaboração entre usuários por meio do compartilhamento dos caminhos realizados durante a navegação e das páginas acessadas, independente de trilha percorrida para atingi-las.

Dentre os trabalhos futuros, visualiza-se a necessidade de aprimorar a modelagem do usuário. Melhorar a eficiência do sistema, uma vez que há uma certa lentidão de acesso. Implementar outras formas de orientação.

6. REFERENCES

- [1] A. Abraham, H. Guo, and H. Liu. *Swarm Intelligent Systems*, chapter Swarm Intelligence: Foundations, Perspectives and Applications, pages 3–25. Springer Verlag, 2006.
- [2] P. Brusilovsky. Methods and techniques of adaptative hypermedia. In *User Modeling and User-Adapted Interaction*, volume 6. Springer Netherlands, July 1996.
- [3] M. Dorigo, V. Maniezzo, and A. Colomi. The Ant System: Optimization by a colony of cooperating

- agents. *IEEE Transactions on Systems, Man, and Cybernetics Part B: Cybernetics*, 26(1):29–41, 1996.
- [4] H. S. Lopes. Fundamentos de computação evolucionária e aplicações. In K. R. Felizardo, editor, *Anais da Escola Regional de Informática da Sociedade Brasileira de Computação*, pages 52–107, Bandeirantes, PR, maio 2006. Departamento de Informática da FFALM.
- [5] L. A. M. Palazzo. Sistemas de hipermídia adaptativa. In *XXI Jornada de Atualização em Informática. XXII Congresso da SBC*, Florianópolis, julho 2002.
- [6] T. Sousa, A. Neves, and A. Silva. Swarm optimisation as a new tool for data mining. In *IPDPS '03: Proceedings of the 17th International Symposium on Parallel and Distributed Processing*, page 144.2, Washington, DC, USA, 2003. IEEE Computer Society.
- [7] P. H. G. Vieira and R. G. Caixeta. Implementação da heurística da formiga para auxiliar à navegação do portal interlegis. 2004.
- [8] A. Zarnani and M. Rahgozar. Mining spatial trends by a colony of cooperative ant agents. In *SIAM Conference on Data Mining*, Bethesda, Maryland, USA, April 2006. Society for Industrial and Applied Mathematics.