Bee Hive Metaphor for Web Search

Pavol Návrat

Abstract: *A new approach to web search that is based on a bee hive metaphor is presented. We proposed a modified model of a bee hive. We have shown that the model is a true model of a bee hive in the sense it simulates several kinds of its typical behaviour. However, more importantly it is a simple model that describes some processes taking place in web search. Having performed several experiments, we have certain confidence the direction of research should be further explored. While we continue performing further experiments, we discuss various possible directions for future research, such as incorporating the concept of bee's memory (including forgetting), or more generally incorporating impreciseness and uncertainty into the model. Other idea is to use the model not only for off line, but also for on line search and subsequent recommendations.*

.
Key words: *Bee Hive Metaphor, Web Search, Information Retrieval, Simulation Model.*

INTRODUCTION

Amount of information and knowledge grows at an unprecedented pace. It may be assumed that most of it is or will be on the web. Therefore, the way information is searched and retrieved becomes crucial.

We propose to explore possibilities open by considering behaviour of social insects. Social insects (ants, honey bees, termites, wasps etc.) show a remarkable level of social behaviour. In particular, they communicate, albeit in a very elementary way, with each other. For example, they may communicate regarding the location of food sources. They even collaborate towards achieving some goal. In particular, they may collaborate regarding bringing the food back once it is found. They distinguish themselves by their organisation skill without any centralised control (Gordon, 1996). The interactions among individuals, between the individuals and the environment along with the behaviours of the individuals themselves allow organisation to emerge.

Our hypothesis is that the behaviour the particular social insects show might be an instructive inspiration to develop a model of searching the web for a source of information. More specifically, we propose to investigate a bee hive as a possibly useful metaphor for web search.

BEE HIVE METAPHOR

Behaviour of honey bees, *Apis mellifera* has been studied extensively by natural scientists, e.g. sociobiologists and behavioral physiologists. One of their aims has been, from our perspective, to formulate a model of bees' behaviour, specifically of their collective nectar source selection. Basically, there are various approaches possible: e.g., to build a simulation model, or an analytical model expressed using differential equations. A mathematical model of (Camazine, Sneyd, 1991) has proven to be especially inspiring. A very elaborated model has later been presented by (de Vries, Biesmeijer, 1998).

Honey bees are social insects who live in colonies. Bees collect food (nectar) from sources distanced up to 10 km from the hive. The colony uses simple rules to dispatch bees towards the best nectar source available. In this process of foraging, the bees return with nectar and the information about its source. Since the sources are not constant, but new ones appear here and there from time to time, as well as the existing ones become exhausted, the information about nectar sources is crucial. The colony keeps on adapting itself to the ever changing situation. To be able to do so, and to maintain a sufficient influx of food, some division of labour force between exploring the surrounding countryside for new sources and exploiting the existing ones is necessary. In other words, there are bees, which go on a foraging mission, but there may be other ones, which are explorers at the given moment.

The foraging bees are able to pass the information they have on the location of the food source they visited onto other bees, and other bees are able to receive that information. Important role in the process of communication plays a waggle dance. Many have studied the mechanism of the waggle dance, the information that is communicated and the way it is actually done (von Frisch, 1967). It is hypothesized that distance and direction are encoded in the waggle dance, although they may not be the only types of information that the foraging bees are guided by. Discussions about these topics, no matter how interesting they may be, are clearly outside the scope of this research. We shall assume that the waggle dance is a means of communicating a food source. The dancer lets know not only distance and direction of the food source; duration of dancing is influenced by the quality of the source.

One interesting experiment aimed at investigating how colonies choose among nectar sources has been devised and performed by (Seeley, 1991). In the vicinity of a colony, there were placed two nectar sources, one of them 400 m to the north, the other one 400 m to the south. 12 bees were trained to fly to the north source, 15 other bees to the south source. The sources were of different quality. The south one was better (sugar concentration of 2.5 units) than the north one (1.0 units) initially. At noon, however, the sources were swapped. Next morning, the sources were swapped again, and at noon once again. Empirical observation showed that number of bees foraging for the better source was increasing in time, whereas number of bees foraging for the worse source remained low. After the quality reversal, the trends reversed, too.

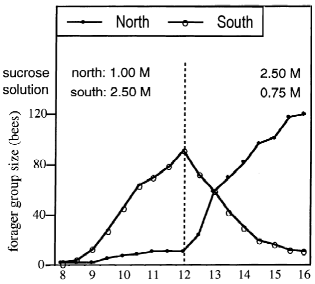

Fig. 1. Result of the experiment aimed at investigating how colonies choose among nectar sources (Seeley, 1991)

Simulation experiments with the model yielded a degree of similarity with the empirical observations.

MODELLING A BEE HIVE

As we could see, behaviour of social insects is interesting in its own right. It deserves a thorough study. For decades, various aspects of their social behaviour have been investigated. Although ants have become perhaps the most popular among them, not only ant colonies, but more recently also bee hives are subject of investigation. Methods of investigation increasingly involve using not only mathematical models, but also computer models. For example, computer simulation models bring new insights into their behaviour. They allow experimenting with various hypetheses.

At the same time, a simple fact was recognized, too: a collective of ants or bees exhibits a large degree of self-organisation, and, indeed, of problem solving capability. This should be studied by sociobiologists, some of them using computer based methods, but perhaps there is some useful inspiration there for those who study methods of computer problem solving, too. Indeed, scientists in fields such as artificial intelligence, information systems and other computing disciplines have started to study behaviour of ant colonies with an aim to improve special problem solving methods of e.g., optimisation, graph searching or document categorization.

Relatively recently, similar expectations regarding a possible model of solving some real world problem have been expressed with respect to honey bee hives. In particular, honey bees behave in a very interesting way achieving remarkable results, viewing them from the problem solving perspective. To consider bees in this context is a relatively new idea (Tovey, 2004; Lorenzi, 2005).

One of the earliest mentions actually dates back to 1986, when Bullock, Dey and Reilly (1986) suggested in a short note a bee hive model for heterogeneous knowledge in expert system, replacing blackboard by a more de-centralised scratchpad. Klugl and Dornhaus (1999) studied methods of multi-agent modeling using as an example a bee recruitment. In a conference poster by Schultze (2002), bee foraging has been applied to enhance the collaborative process of filtering information.

One of the most recent projects by Lorenzi et al. (2005) makes use of the bee hive metaphor in a case-based recommender system. Their idea is to use bee dance to retrieve the most similar case to the user's query. They adopted model of Camazine and Sneyd (1991) and combined it with case-base reasoning. Case corresponds to a nectar source. Case base is a set of nectar sources. In a standard case-based recommender, a user query is compared with all cases in the case base. The most similar is returned to the user. In their recommender, bees decide on abandoning or continuing to visit the case depending how similar is the query to the visited case. In other words, the probability of abandoning the source decreases with greater similarity between the query and the case. Bees start foraging randomly and after some time, a case with the clearly highest number of bees emerges.

Let us now proceed to presenting our model of a colony of bees living in a hive. Its rationale is the following. We want to explore possibilities of defining the process of searching, retrieving etc. information in a way that is similar to the process of searching, bringing etc. food to the bee hive. Intentionally, this is a rather broad formulation, since we want to research in various directions, not making assumptions from the start what will be the best one, if any. First thing to do was to devise a model of a bee hive, that would both be consistent with the biological original and suitable for experimenting in information retrieval.

Therefore, there are two aims regarding our model: it should model behaviour of biological bee hive and it should model information retrieval. However, our main objective is a possible contribution to information retrieval. We shall devise a bee hive model, but we do not obviously aim at a contribution to research in biology. It will suffice if our model exhibits a degree of similarity with the bees' foraging behaviour. One of the properties of our model should be simplicity.

We took an initial inspiration from the model of Lorenzi et al. (2005) who in turn were inspired by Camazine and Sneyd (1991). Our model of the bee hive comprises of a dance floor, auditorium, and dispatch room. We enriched their model by introducing a dispatch room.

The dispatch room brings additional flexibility to the model. It is an explicit reflection of the hypothesis that some bees in the auditorium may not be ready to follow any one of the advertised sources. They may either not become enthusiastic for any dancer within some reasonable time, or there may not be any dancers to observe within some reasonable time. Therefore, they simply decide better to go foraging blindly than run into a danger of starving in the auditorium. Moreover, bees take off from the dispatch room initially, when a new query arrives.

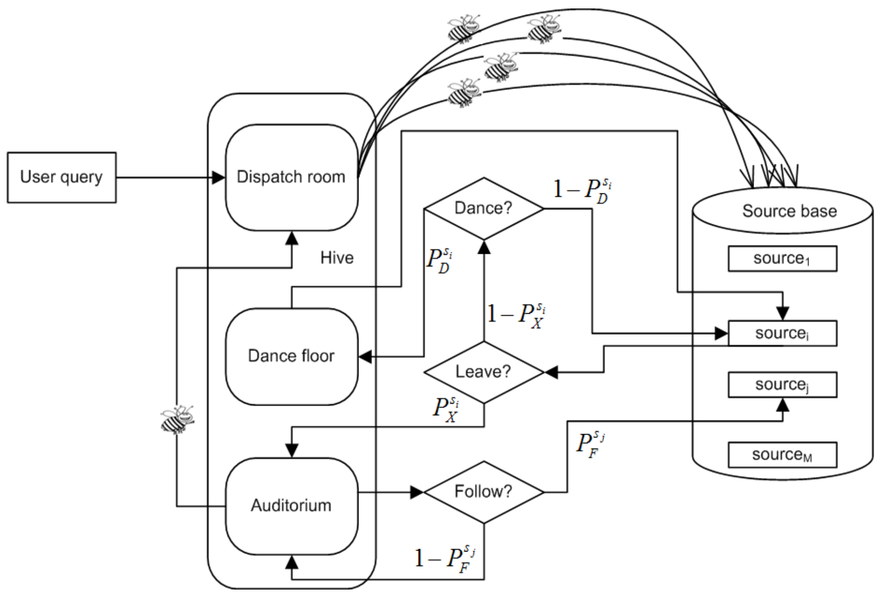

Fig. 2. The bee hive comprises of the *dance floor*, *auditorium* and *dispatch room*

The model is flexible due to several its attributes which are parametrised. Parameters of our model are the following:

- overall number of bees – in the initial state, it consists of bees in the auditorium, and those which start foraging,
- initial distribution of bees between observers and foragers,
- maximal time allotted for dancing for a source (maximal dancing time),
- maximal time a bee can spend in the auditorium (observing time).

It should be noted that the time a particular bee dances depends on a quality of the source (it equals maximal dancing time*quality). The observing time, on the other hand, is either the maximal one or is shorter if the particular bee was attracted by some dancer before the maximal observing time elapsed.

The process of searching for the best source starts when a user query arrives. In response to that, forager bees take off from the dispatch room randomly, i.e. each of them is sent to a randomly chosen source. Each bee, having visited the source, evaluates it with respect to the user query. The quality Q of a source is defined as a measure how much the source matches the query.

Having returned to the hive, the bee decides if she abandons the source she just visited (depicted as the selection diamond *Leave?*). The model defines a probability $P_X^{s_i}$ ($P_X^{s_i} = 1 - Q$) that the bee abandons the source. The probability of not abandoning the source is obviously $1 - P_X^{s_i}$. If the bee decides to abandon the source, she immediately enters the auditorium.

If the bee stays with the source, she has again two options. She can either try to attract some fellow bees for the source or she can take off for another visit to the source (depicted as the selection diamond *Dance?*). The model defines a probability $P_D^{s_i}$ ($P_D^{s_i} = 1 - P_X^{s_i}$) that the bee goes dancing for the source. The probability of not going to dance is obviously $1 - P_D^{s_i}$. In that case she flies to the source. If the bee decides to dance for a source, she immediately enters the dance floor. The better the source quality, the longer she dances for her source. However, the dancing time cannot be longer than MDT.

After the bee completes her dancing engagement, she flies to the source which she was trying to advetise by dancing.
Bees in auditorium observe dancing bees. They try to make a decision which dancer to follow for a source. An observing bee chooses a source by choosing randomly some dancer. Let us assume this dancer advertises a source s_j. She will fly to this source with a probability $P_F^{s_j}$. The model defines the probability $P_F^{s_j}$ as the number of bees dancing for the source s_j divided by the number of all dancers. The probability that the bee will not fly to the source s_j is $1 - P_F^{s_j}$. In that case, the bee will choose randomly another dancer.
If the bee does not choose a source within the alloted time, she moves from the auditorium to the dispatch room. From it, she flies to a randomly assigned source.
The whole process repeats as described here.
The model just described here is a simplification of models studied by biologists. Still, we hypothesize that it is, at least in some respect, a model of a bee hive. So we do not attempt to model all the aspects of a bee hive's behaviour, having in mind our main goal of developing a model of web search. We tried to perform a similar experiment as the one published by (Seeley, 1991), because this experiment has been used in the literature as a kind of testbed for bee hive models. Results of our experiments show that our model conforms with the expectations of the original experiment.

MODELLING INFORMATION RETRIEVAL

We hypothesise that the bee hive metaphor could prove useful in devising methods of web search, and in particular, of information retrieval from web. We investigate several possible directions of research, including a possible contribution to information recommendation or filtering. At this stage, we wish to report on experiments that may shed some additional light into these considerations, stimulating at the same time further discussion.
Documents on the web have usually several attributes. Some of them can have only one out of two values, some can have value from an interval of real numbers. This is a fairly general assumption, although we definitely do not claim it covers all possible cases. Let us assume, for the simplicity sake, that the binary valued attributes can be either 0 or 1 and the value 1 is the one matching the user query. Let us further assume that the real valued attributes can have value from the interval [0,1] and higher value better matches the user query.
We have performed a series of experiments, out of which we shall report here only two. In both experiments, we assumed that the sources have three attributes. Two of them are binary and one is real. In the first experiment, a bee was able to evaluate the source how close is it to the user query by considering all the three attributes. In the second experiment, we tried create a situation when some bees will during foraging use only the first attribute, some will use only the second attribute and some will use only the third attribute.
In the second experiment, some bees contribute to recommendations by evaluating only the first attribute, some only the second attribute and some only the third attribute. A source with better evaluations in each attribute will be rewarded by higher number of combined bee visits. Objective of our experiment was to see if this hypothesis deserves further attention or it should be rejected.
We are in the process of experimenting. First results seem to indicate that the recommending of retrieved information based on our bee hive model works in both cases. Most of the bees decide to go foraging for the best source.

CONCLUSIONS AND FUTURE WORK

A new approach to web search that is based on a bee hive metaphor was presented. Building on related works especially by Lorenzi et al. (2005), we proposed a modified model of a bee hive. We have shown that the model is a true model of a bee hive in the sense it simulates several kinds of its typical behaviour. However, we have been even more keen to formulate a simple model that would describe some processes taking place in web search. We claim we have such a model, as some of our experiments show.

We are well aware most of the road in the research is still ahead. The bee hive metaphor has received further endorsement. At the same time, to reach a workable model of web search requires not only refinement, but more ideas. The idea of bee typing according to types of attributes they evaluate is to be elaborated. We have not attempted to introduce any schemas how to combine partial results, although there are obvious candidates such as weighting etc. The model itself can be more refined. Additional parameters, such as initial distribution of foraging directions could be considered. Alternatively to generating initial direction randomly, some kind of memory (individual or collective) could be considered helping to retrieve directions to sources which have been useful in the past in providing information relevant to the current or similar past query.

Some reflection on the very purpose of the model might be desirable, too. We have used the term web search in a rather loose sense. It appears as retrieving information from the web could profit from this approach. In particular, information recommending and filtering could be modelled by the bees' behaviour. But other possibilities are still open for research.

The bee hive metaphor has been used so far as a basis for a model of recommending information from a given set of sources collected from the web. Besides such an off line application, searching a "frozen" part of the web, it might be worthwhile to attempt to search the "living" web with bees becoming agents. Such a dynamic search, if performed continuously, could open possibility of adaptive recommendation according to changes (disappearances of old documents and appearances of new ones) on the web.

ACKNOWLEDGEMENTS

This work was partially supported by the Slovak State Programme of Research and Development "Establishing of Information Society" under the contract No. 1025/04 and the Scientific Grant Agency of Republic of Slovakia, grant No. VG1/3102/06.

REFERENCES

[1] Bullock, H.E., Dey, P., Reilly, K.D.: A "Bee Hive" Model for Heterogeneous Knowedge in Expert Systems. ACM, 1986, 417-417.

[2] Camazine, S., Sneyd, J.: A Model of Collective Nectar Source Selection by Honey Bees: Self Organization through Simple Rules. Journal of Theoretical Biology, 1991, 149(4): 547-571.

[3] de Vries, H., Biesmeijer, J.C.: Modelling Collective Foraging by Means of Individual Behaviour Rules in Honey-Bees, Behav Ecol Sociobiol (1998) 109-124.

[4] Drogoul A. and Ferber J.: Multi-Agent Simulation as a Tool for Modeling Societies: Application to Social Differentiation in Ant Colonies, C. Castelfranchi and E. Werner (eds.). Artificial Social Systems (MAAMAW92), Springer 1992, pp. 3-23.

[5] Gordon, D.: The Organization of Work in Social Insect Colonies. Nature, 1996, 380, 121-124.

[6] Kluegl, F., Oechslain, C., Puppe, F., Dornhaus, A.: Multi-Agent Modelling in Comparison to Standard Modelling. In: Barros, F.J., Giambasi, N. (Eds.): Proc. AIS2002 Artificial Intelligence, Simpulation and Planning in High Autonomy Systems, SCS Publ. House (2002) 105-110.

[7] Lorenzi, F., Scherer dos Santos, D., Bazzan, A.L.C.: Negotiation fopr Task Allocation Among Agents in Case/Base Recommender Systems: A Swarm Intelligence Approach. In: Proc. IJCAI 2005 Conference, Workshop, 2005, 23-27.
[8] Lorenzi, F., Scherer dos Santos, D., Bazzan, A.L.C.: Case-Based Recommender System Inspired by Social Insects. In: Proc. XXV Congresso da Sociedade Braseilera de Computacao, Sao Leopoldo, 2005, 752-760.
[9] Schultze, S.J.: A Collaborative Foraging Approach to Web Browsing Enrichment. In: Proc. CHI 2002, ACM, 2002, 860-861.
[10] Seeley, T.D., Camazine, S., Sneyd, J.: Collective Decision-Making in Honey Bees: How Colonies Choose Among Nectar Sources. Behav Ecol Sociobiol 1991, 28, 277-290.
[11] Tovey, C.A.: The Honey Bee Algorithm. A Biologically Inspired Approach to Internet Server Optimization. Engineering Enterprise, Spring 2004, 13-15.
[12] Visscher, K., Camazine, S.: Collective Decisions and Cognition in Bees. Nature, 1999, 397, 400-400.

ABOUT THE AUTHOR

Prof. Pavol Návrat, PhD. navrat@fiit.stuba.sk, from:

Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies, Slovak University of Technology in Bratislava, Slovakia.