Community-Aware Mobile Networking

Michael Kuhn
Computer Engineering and
Networks Laboratory
ETH Zurich
8092 Zurich, Switzerland
kuhnmi@tik.ee.ethz.ch

Roger Wattenhofer
Computer Engineering and
Networks Laboratory
ETH Zurich
8092 Zurich, Switzerland
wattenhofer@tik.ee.ethz.ch

Abstract:

Community related applications such as dating platforms, instant messengers, or file-sharing tools enjoy a great popularity in the Internet. Motivated by this success, community-awareness also has become a hot topic in the mobile world. Despite considerable efforts, however, mobile solutions still cannot compete with their Internet counterparts. A key feature of community-aware applications is profile matching. To a large extent, the missing popularity of mobile approaches can be explained with the problems traditional one-to-one keyword matching causes in a purely decentralized environment: Expressive keywords are unlikely to match because of a lack in user density, and a superficial keyword classification misses the application needs. Hence, we face the challenge to compare characteristic topics in purely mobile applications. To address this challenge, this paper presents a fundamentally new matching scheme based on a combination of topology analysis and distance labeling. Proof-of-concept experiments demonstrate that the proposed scheme has indeed the potential to facilitate the comparison of highly specialized keywords in mobile environments.

1 Introduction

By means of e-mail, news groups, and instant messengers, the Internet has crucially altered the way people communicate. Community related applications, such as Friendster (`www.friendster.com`), ICQ (`www.icq.com`), dating services, or file-sharing platforms have become extremely popular. Not surprisingly, similar applications have started to enter the ever growing market of mobile devices. More surprisingly, so far none of these mobile applications managed to follow the success of their wired counterparts.

One reason for the missing popularity of community-aware mobile applications is that they often rely on one-to-one keyword matching of metadata and user profiles. Such a simple strategy works well in the Internet where a match is likely to be found among the huge amount of profiles stored on a central server, but it is not well suited for the mobile world. Either the mobility is compromised by contacting some central infrastructure—which is hardly ever free of charge—or the information needs to be extremely superficial in order to find an exact match in the vicinity of the device.

We believe that community-supporting mobile applications can become extremely successful, provided they are independent on central infrastructure, free of charge, reflect the

underlying topics equally detailed as the wired competitors and fit the usage pattern of mobile devices, which in contrast to the wired world is not place and time, but event and person based.

We show that a combination of topology analysis and distance labeling allows a fully decentralized comparison of data that is equally fine grained as in typical Internet applications. Throughout this paper, we will mostly focus on two applications that doubtlessly have an enormous potential: Friend-finding (including dating) and file-sharing.

The rest of this paper is organized as follows. Section 2 provides some background about the history of communities and how mobile networking could affect their future. Section 3 then explains how mobile devices can become community-aware, followed by a demonstration of the feasibility of the developed concept in Sect. 4. The paper concludes with a discussion of the involved challenges and related work in Sects. 5 and 6, respectively.

2 Background

In the very beginning, the main function of social groups presumably was the increased chance of survival. Today, the driving forces are having fun in common leisure activities, setting up business contacts, or striving for romantic encounters or an inner sense of well-being. Already with the beginning of writing, the invention of the wheel, and the domestication of the horse, the bonds of locality started to disrupt and allowed longer distance relationships to appear. Together with trains, cars and airplanes, information technology in the shape of newspapers, television and the Internet transformed huge geographical distances into tiny fragments of time. In 1964, Marshall McLuhan coined the metaphor "Global Village" and wrote:

"As electrically contracted, the globe is no more than a village" [McL64].

He argues that due to the almost instantaneous reaction times of new ("electric") technologies, each individual inevitably feels the consequences of his actions and thus automatically deeply participates in the global society. McLuhan understood what we now can directly observe—real and virtual world are moving together.

Nowadays Internet communities are widely accepted. Recent studies show that online communication adds to rather than replaces face-to-face meetings, and that it fosters new contacts not only in the virtual, but also in the real world [TSB98, Wel05]. However, some major issues remain with the Internet approach. First of all, people that cannot frequently access the Internet are in danger to miss contact to the future society. Further, a trend from group to bilateral, and from person to role based interaction can be observed. We have more friends, but know less friends of friends [DHJ$^+$02, Wel05], and for every single need, such as material aid, a piece of information, or emotional support, we quickly connect to a new node in the virtual community. The consequence are worries that the individual degenerates to the sum of its roles and thus loses its identity [Wel00].

We are convinced that mobile devices are the perfect gateways between real- and cyberspace. Their high availability[1] ensures that almost everybody can populate the new square that connects cyberstreets with concrete roads. When evolving to mobile agents, they further have the potential to combine place with person as well as person with role based interaction.

Imagine that a friend you have not seen for a long time, or maybe the perfect match for you lonely heart is just around the corner. Wouldn't it be great to have an agent that informs you about such a "neighbor in the global village"? You could start a face-to-face conversation that allows you to get to know someone with an own identity and not just some node representing a role in the networked society. The device could inform you about a friend of a friend in your surrounding or directly introduce you to the vendor of exactly that vintage car you have been looking for for years.

Human interaction, however, should be restricted to those situations that really ask for it. Your mp3-player, for example, does not need to disturb you in order to exchange files with someone who has songs of your taste on his device. A mobile agent should extend our awareness of the social environment in a natural way. Nobody is willing to conduct a lengthy conversation using the clumsy keyboard of a cell phone, or to browse a whole lot of information on the tiny display. Instead, the device should act as a natural companion that mostly works autonomously and unnoticeably gives notice about important events if necessary.

3   Mobile Keyword Matching

Topic matching is doubtlessly crucial to community aware applications, which typically need to compare people or object profiles[2]. Traditional Internet solutions often rely on direct keyword matching. Unfortunately, simple one-to-one matching only works if either the degree of detail is low, or the number of profiles is high, which is usually not the case in mobile environments. Neither one-to-one matching nor generalization could, for example, appropriately reflect both, the close relations from "Aston Martin DB5" to "Ferrari 275" (both sport cars from the 1960's) and to the movie "Goldfinger" (in which James Bond drove a DB5). Different tricks try to address these issues in today's mobile solutions (see Sect. 6.1), but none of them attacks the heart of the problem: Matching expressive keywords in a purely decentralized environment.

The relations stated in the above example about the "Aston Martin DB5" are well reflected in the link structure of thematic graphs, such as an encyclopedia, or the World Wide Web: An article treating the "DB5" is likely to contain links to articles about James Bond as well as ancient sport cars. The analysis of such link structures has proven to be a powerful technique to infer topic similarities (see Sect. 6.2). Most work in the area focuses on the hyperlink structure of the World Wide Web. However, topology based methods have also been applied to scientific publications [Gar72, Sma74], US patents [Kle98] or music

---

[1]It is said that everybody earning more than 5 USD a day will eventually possess his/her own mobile phone.
[2]The latter type of "profile" is also referred to as metadata.

graphs [GRZL05, Pla04].

Directly using topology analysis in the mobile world is difficult, since access to the Internet (or a database containing the required graph) is costly, and the devices lack computing power and memory to deal with the immense data amounts involved. To nevertheless exploit the power of topology analysis we propose to assign distance labels in a preprocessing step. Distance labels are encodings assigned to each node of a graph. These labels allow to derive the distance between any two nodes without having to know and store the structure of the underlying graph. That is, the distance between two nodes is given as a function of the two labels: $d(v_1, v_2) = f(l_{v_1}, l_{v_2})$, where l_v denotes the label of node v. In our case, these distances do not necessarily need to reflect shortest paths, but rather some sort of relatedness of the nodes within the graph.

Figure 1 illustrates the use of such encodings for the example of a mobile file-sharing application. First, distance labels are assigned to each node of a graph that stores information about artists and songs, such as Wikipedia (`www.wikipedia.org`) or AllMusic (`www.allmusic.com`). The metadata (i.e artist and song) together with its distance label is then made available through a central server (1).[3] Whenever a user wants to upload a song to the mp3-player, he/she does so by using a special software. This software acquires the proper distance label (based on the title, etc.) from the server (2). The file is then transferred to the device together with the label (3). The device can now be disconnected from the PC, and the system becomes entirely independent from any fixed infrastructure. Whenever two devices come into connection range (through Bluetooth, for example) (4), they can derive the similarity of the stored music based on the labels of their files (similar labels result in short distances and thus refer to closely related songs) (5). The songs on the remote device that match the taste of the user can then automatically be downloaded, together with their labels (6).
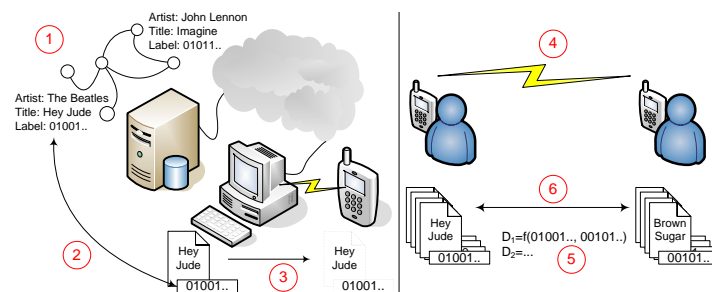

Figure 1: Distance labels in a mobile file-sharing application. Left: Uploading files to the mobile device. Right: Sharing files among different users.

The same principle can of course not only be used for the comparison of music files, but for any kind of topic, such as for keywords defining a user profile in a matchmaking application, for example.

---

[3]This step (assigning labels and publishing them) has to be repeated once in a while, to account for newly released songs.

4 Experiments

This section presents some experiments that demonstrate the creation of labels for the scheme introduced in Sect. 3. In these experiments, we concentrated on the graph of Wikipedia. A key advantage of using an encyclopedia is that each topic is directly defined by an article's title, which minimizes the effort of topic identification. Hyperlinks to other entries define the structure of the graph. There is a high density of links within groups of closely related topics, whereas only few links exist between uncorrelated areas, similar as in the case of the World Wide Web.

In the following discussion we assume that a graph $G(V, E)$ has $|V| = N$ nodes and $|E| = M$ edges.

4.1 Shortest Path Analysis

Path length calculation provides a simple approximation for thematic distances within the Wikipedia graph. There are some inherent problems to directly using shortest paths, however. Due to the small-world character[4] of the Wikipedia graph [Vos05], some articles exhibit a very high in-degree, and consequently act as a hub coupling many unrelated topics. In addition, some links do not well reflect thematic relations, such as a link from "cow" to "Bible", for example.

To some extent, the first problem can be addressed by assigning degree dependent edge weights to the graph. The higher a node's degree, the less relevant a path through it becomes. For our experiments, an edge between node v_i and v_j was symmetrically weighted with

$$w_{ij} = w_{ji} = \left[\frac{1}{c} \sqrt{\deg(v_i) + \deg(v_j)} \right],$$

where $c = 2$ if the link is bidirectional in the original graph, and $c = 1$ otherwise. Table 1 shows the closest nodes to the topics "Internet" and "George W. Bush" within the simple English Wikipedia[5]. Despite of a clear outlier ("Llanfarip..."[6]), these lists show that the shortest path approximation is a good starting point. Unfortunately, the time complexity for all-pairs shortest paths on sparse graphs is $O(N^2 \log N + NM)$, which, together with the memory requirements of $O(N^2)$ to only save the results, makes it impractical for large graphs such as the (normal) English Wikipedia with its approximately one million articles.

Another problem to directly using shortest paths would be to find an appropriate distance labeling scheme. Lower bounds on the number of bits per label are shown to be $\Omega(\sqrt{N})$

---

[4]Small world character means that any two nodes are separated by few hops only, and that the vertex degrees follow a power law distribution.

[5]The simple English Wikipedia (as opposed to the (normal) English Wikipedia) was used, since meaningful lists of topics become shorter and the experiments ask for less computing power and memory resources.

[6]The place with the longest name in Britain.

Table 1: The 10 closest topics to "Internet" and "George W. Bush" together with the corresponding distances (D).

Internet	D	George W. Bush	D
Computer network	5	Laura Bush	4
World Wide Web	6	Barbara Bush	4
Instant message	9	Jenna Bush	4
Virtual community	9	Dick Cheney	5
Computer	9	July 6	5
LiveJournal	10	George H.W. Bush	5
Llanfairpwllgwy...	10	John Kerry	5
Chat room	10	Barbara P. Bush	6
Local area network	10	Texas	6
Html	10	Vice President	7

for bounded degree graphs; For general graphs even $\Omega(N)$ bits are required for any approximation with a stretch $s < 3$ [GPPR04].

4.2 Graph Embeddings

A method that has the potential to overcome the outlier as well as the labeling problem, and at the same time takes advantage of the quality and simplicity of the shortest path approach, is graph embedding. The goal is to map the high dimensional graph structure onto a lower dimensional euclidean space while approximately preserving distances.[7] Figure 2 shows an embedding of the distances between 5 US cities (see Table 2) in 2 dimensions.

The euclidean coordinates resulting from an embedding can directly be used as the distance labels for the respective nodes and outliers are likely to be a minor problem in this approach, since a large amount of other paths can compensate for a single short edge.

Table 2: Distance matrix between 5 major US cities.

	CHI	LA	MIA	NY	SF
CHI	0	2054	1329	802	2142
LA	2054	0	2687	2786	379
MIA	1329	2697	0	1308	3053
NY	802	2786	1308	0	2934
SF	2142	379	3053	2934	0

Classical multidimensional scaling (MDS[8]) is one of the most prevalent dimension reduc-

---

[7]Bourgain [Bou85] showed, that an embedding with $O(\log N)$ distortion is always possible into an euclidean space with $O(\log^2 N)$ dimensions.

[8]Although MDS usually refers to any sort of multidimensional scaling, classical multidimensional scaling is

Miami
(-1227/1014)

Los Angeles
(1464/561)

San Francisco
(1697/132)

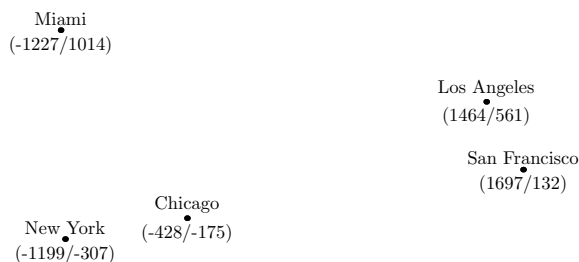Chicago
(-428/-175)

New York
(-1199/-307)

Figure 2: A 2-dimensional embedding of the (complete) graph resulting from the distance matrix given in Table 2. The embedding only defines the relative positions among the nodes but not the direction of the coordinate axes.

tion techniques. It minimizes the sum of the squared errors between the reconstructed and the observed distances. That is, among all the possible projections into an euclidean space of given dimension, the one that minimizes $\sum_{i=1}^{N} \sum_{j>i}^{N} (\delta_{ij} - d_{ij})^2$ is found, where δ_{ij} are the observed, and d_{ij} the reconstructed distances. The complete distance matrix is required at the input and an eigenvalue problem is solved to find the locations that minimize the distortions. Consequently, all-pairs shortest paths have to be calculated in order to apply MDS, which is impractical as stated before. To cope with this problem, de Silva and Tenenbaum [dST02, dST04] suggested the landmark MDS (LMDS) algorithm, which applies MDS only on a subset of l landmark nodes and places the remaining nodes according to their distances from these landmarks. LMDS asks for l single source shortest path evaluations only and has been successfully applied before [Pla04].

4.3 Experiments with LMDS

This section summarizes the major findings of the use of LMDS in conjunction with Wikipedia.

The impact of the number of output dimensions on the embedding quality is illustrated in Table 3, which compares the closest topics to "The Beatles" in a 2D and a 30D LMDS embedding of the edge weighted simple English Wikipedia graph. This comparison shows that neighboring nodes in the 2D embedding have rather random character, while those in the 30D case have a good thematic justification.[9] A closer look reveals that in 2D most nodes do not even lie in the 2-hop neighborhood of the "Beatles"-node, but are there due to a lack in space, which is reflected by the short distances stated in Table 3.

A comparison between the quality of the results of the LMDS embedding and those of the shortest path analysis (recall Table 1) is given in Table 4 for the topic "Internet". Besides the disappearance of the outlier ("Llanfairp...") in the embedding, there are only minor differences and none of the methods clearly outperforms the other one.

---

meant throughout this paper.

[9]All of the 10 topics are in the 2-hop neighborhood of "The Beatles".

Table 3: Left: The 10 closest topics to "The Beatles" for a 2- and a 30-dimensional LMDS embedding of the simple English Wikipedia together with the distances (D). Right: A close-up of the area around the "Beatles" node that illustrates the lack in space in the 2D LMDS embedding. The only two edges that lie entirely in the selected area are highlighted black.

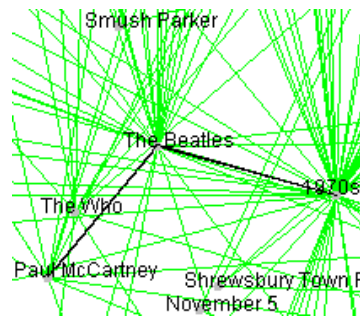| 2D	D	30D	D
The Who	0.11	Mersey Beat	3.91
Smush Parker	0.14	Paul McCartney	4.05
Shrewsbury T. F.C.	0.17	Ringo Starr	4.98
Paul McCartney	0.19	Liverpool	5.32
1970s	0.19	George Harrison	5.61
November 5	0.19	John Lennon	5.95
December 3	0.24	December 8	6.29
Kiddermin. H. F.C.	0.24	Elvis Costello	6.38
York City F.C.	0.24	July 7	6.74
Barnet F.C.	0.24	Badfinger	6.96

Table 4: The ten closest topics to "Internet" in the simple English Wikipedia for an LMDS embedding and a shortest path analysis together with the distances (D).

LMDS 30D	D	Shortest Path	D
Instant message	4.76	Computer network	5
Virtual community	5.23	Word Wide Web	6
Computer network	5.49	Instant message	9
World Wide Web	6.88	Virtual community	9
Html	7.41	Computer	9
Chat room	7.49	LiveJournal	10
Computer	7.81	Llanfairpwllgwy...	10
IP address	7.81	Chat room	10
Computer jargon	7.99	Local area network	10
LiveJournal	8.41	Html	10

The results presented in this section show that Wikipedia features a significant topical clustering that can be extracted and made available to the mobile world by embedding or—presumably—some other labeling technique.

5   Design Issues

5.1   Distance Labeling

Despite of the existing research about quantifying topic similarities from link topologies, the problem of generating labels that appropriately reflect these similarities raises several

design issues.

This paper provides one possible way (high dimensional graph embedding) to approach the labeling problem. Our experiments with a "general-purpose" algorithm (LMDS) show the high potential of this technique. However, tailoring embedding algorithms at exactly this application would facilitate significantly better results. Such algorithms should optimally reflect the clustering characteristics of graphs, put less stress on distortions with respect to shortest paths, and, unlike MDS, provide a high accuracy particularly for short distances[10].

A fundamentally different approach is the development of a more abstract relationship labeling scheme. It could stick closer to the graph structure and not only rely on shortest paths, but, for example, base on a maximum flow analysis and thus allow to reflect the desired clustering properties more appropriately.

Further, ideas from both directions could be combined. A preprocessing step based on topology analysis could introduce new or eliminate existing edges and assign more appropriate weights. An off-the-shelf embedding algorithm might then produce an accurate thematic labeling from this modified graph.

Further, better sources than Wikipedia are likely to exist, such as other encyclopedias or the World Wide Web. To cover specialized subject areas more precisely, small and detailed, thematically bounded graphs could be applied, or subsets of larger graphs could be extracted and analyzed in depth.

5.2 Dynamics

In our approach, the labels are assigned once, before they are made available through a central server. To keep the system up-to-date, this assignment has to be repeated periodically. This process, however, has to make sure that the existing labels do not (significantly) change, such that they are still appropriate. Consequently, a labeling scheme should allow to add new nodes without affecting the existing labels. In the LMDS approach, for example, new nodes could be placed based on their distances to the landmark nodes, neglecting their influence on other (existing) paths.

5.3 Topic Extraction

Manually entering keywords into a profile is simple to implement, but exhibits several drawbacks. Once entered, the information will rarely be adjusted and does therefore not adequately reflect a user's current activities. Further, a high degree of self-awareness is required to appropriately select topics. Finally, each user has his own methodology of selecting topics, resulting in different patterns that are difficult to compare.

A more promising technique is the extensive collection of usage information, not only on the mobile device, but when- and wherever possible. Yenta [Fon97], an agent based

---

[10]We are not interested in how unrelated unrelated topics are.

online matchmaking system, for example, extracts the necessary information by textual inspection of emails, newsgroup messages and user files. Other usage data, such as the surfing behavior could also be taken into account. If, for example, a person mostly opens threads in a forum, he/she is likely to be a novice in the respective field, while those answering are likely to be experts. Similarly, a file-sharing application might profit from usage statistics of the respective files. Information collected by the social environment, or the investigation of success and failure rates could finally be used for fine-tuning.

Automatic profile generation, possibly combined with manually entered keywords to fill gaps, hopefully enables the design of dynamic profiles that facilitate mobile agents that flirt with one-another rather than just proclaiming lifeless facts.

5.4 Social-Awareness

In real world many new acquaintances are doubtlessly facilitated through existing contacts. People get redirected to a friend's friend when asking for some expert knowledge, for example. The success of Friendster and MySpace (`www.myspace.com`), both online friend finder systems based on social networking, suggests to consider such information in other applications, too.

Mobile phones contain a considerable amount of information about our social networks in the form of the contact book. This information can easily be exploited in a purely decentralized environment. When two devices come into connection range, the one-way hashes of the saved phone numbers[11] can be compared. Two matching entries identify a common friend and consequently the device owners are separated by only two hops in their social networks. Only users that do not yet know each other—and thus do not have stored each others numbers directly—have to be informed about common friends, of course.

Obviously, the system could be significantly improved by considering the usage frequency of numbers. Once the importance of a contact is quantified, variations, such as caching of a close friend's most important numbers in order to detect high quality nodes in the three- or even four-hop-environment, become possible.[12]

5.5 Contextual Cues

A comprehensive system should consider more than just topic similarities. A predestined cue to be taken into account in mobile systems is location. A person sitting in the same subway as you every day is of little interest when sitting in the subway. When, however, meeting exactly the same person somewhere in the middle of Siberia, there is a high probability of a mutual interest to get acquainted.

---

[11]And e-mail addresses, ICQ numbers, etc.

[12]Note that considering more than four hops is not likely to be interesting when having Milgram's six degree of separation theory in mind [Mil67].

A similar effect applies to time. An animated talk might result when your mobile phone can tell you that the person next to you in the train went to the same school some ten years ago. The effect is even amplified when playing together with its location pendant.

Concise facts, such as languages spoken, nationality or the age might further help to improve the success rate. This data can be entered manually, and, once entered, automatically be weighted and adapted based on the device's experience.

Contextual cues can and should be used for both, to foster new social interaction, but also to avoid undesired disturbances.

5.6   Privacy Issues

The closer on- and offline lives move together, the more important the boundaries of in-person interactions become in the world of cybercommunities. Mobile agents should have their rules of conduct, they should, for example, not provide considerably more information about their owner, than they in turn learn about their conversation partner [Fon97]. If our good-natured mobile companions naively disclose whatever they know, they could not only greatly embarrass their owner, but might also get abused seriously.

Applications face the challenge to compare distance labels or phone numbers and at the same time reveal as little information as possible. Distance labels that do not lead to a match should contain as few information as possible about a user's interests, and similarly, only common friends should be known after comparing contacts, even when examining the four-hop neighborhood.

6   Related Work

6.1   Community Aware Mobile Systems

Many socializing applications and devices have been proposed for the mobile world. Among them a variety has the objective of bringing people together, such as Lovegety [Iwa98], FriendZone [BS04], Bedd (`www.bedd.com`), MobiLuck (`www.mobiluck.com`), Jambo (`www.jambo.net`), 6th Sense (`www.sixsense.com`) or Nokia Sensor (`www.nokia.com/sensor`). All these systems rely on profile matching, but differ in the way they tackle the problem of expressiveness and user density. Lovegety makes use of very general profiles (only 3 bits are required to store the possible states) to increase the chance of a match. Other applications artificially increase the user density by accessing the Internet (6th Sense, Jambo, MobiLuck, FriendZone), using a store-and-forward technique (Bedd), or introducing meeting-points (MobiLuck). Finally, the problem is addressed at the matching function itself, either by relying on the user doing the matching (Nokia Sensor), or by transferring the task to a powerful server with access to huge databases (6th Sense).

Other, more inventive systems that aim at bringing people together exist, such as Serendipity [EP05], Just-for-us [KP05], or AgentSalon [SM00]. Serendipity closely follows the matchmaking idea. Through a central server, information from online services, such as Friendster, Monster (`www.monster.com`) or Match.com (`www.match.com`) are accessed, and people within the Bluetooth communication range are alerted in case of a mutual match. The authors emphasize that the system is not only headed at dating, but might also prove helpful in more general situations such as the working place, in conferences or for car-sharing, for example. Just-for-us also bases on a central database that stores information about the location, such as entertainment places nearby, as well as about friends and their whereabouts. It is not mainly designed for new contacts, but to enhance sociality among existing friends. AgentSalon finally tries to facilitate face-to-face contacts between people that visit some common event, such as an exhibition. A large-scale display with animated characters stimulates the conversation between the participating users.

But not all systems aim at matchmaking or bringing people together, some also make use of other community related features. Jabberwocky [PG04] and Telelogs [DK05] take advantage of the familiar stranger phenomena studied by Milgram, which expresses the fact that many people regularly see but nevertheless ignore each other.

According to Motani et al. [MSN05], humans are the best source of information, particularly in time, location and community specific settings. Their system PeopleNet propagates question and answer queries in a word-of-mouth manner through so called bazaars, which are thematic clusters of users. While the communication inside a bazaar is Bluetooth based, the bazaars itself are reached by the network infrastructure, which is also used to inform a user of a successful match.

ContextContacts [ORT05] finally aims at lowering the disturbance of mobile phones by adding cues about the callee's location, the time spent there and the active alarm profile to the contact book. The cues are distributed by SMS, which makes the system rather costly.

6.2 Algorithms

Inferring topic similarities from link structures is a well-known approach. The arguably best studied algorithm in the area is Hyperlink Induced Topic Search (HITS) [Kle98], which recursively adjusts node weights based on incoming and outgoing links, similar to the (published) PageRank algorithm [PBMW98] used by Google. Other approaches try to estimate the maximum flow between different nodes [FLGC02], or apply some sort of embedding algorithm [GRZL05, Pla04].

Gavoille et al. [GPPR04] have proven lower bounds for exact distance labeling for various classes of graphs. These bounds show that exact schemes are unlikely to solve our problem. The results concerning approximate distance labeling are more promising. A 1-additive scheme that requires $O(\log^2 N)$ bits per label and $O(1)$ decoding time exists, for example, for chordal N-node graphs [GKK$^+$00].

References

[Bou85]   Jean Bourgain. On Lipschitz embeddings of finite metric spaces in Hilbert space. *Israel Journal of Mathematics*, 52(1-2):46–52, 1985.

[BS04]   Asaf Burak and Taly Sharon. Usage Patterns in FriendZone - Mobile Location-Based Community Services. In *Proceedings of the 3rd International Conference on Mobile and Ubiquitious Multimedia*, pages 93–100, 2004.

[DHJ$^+$02]   Tom Denison, Gary Hardy, Graeme Johanson, Larry Stillman, and Don Schauder. Community Networks: Identities, Taxonomies and Evaluations. In *Electronic Networking - Building Community Conference*, 2002.

[DK05]   Brian Davis and Karrie Karahalios. Telelogs: A Social Communication Space for Urban Environments. In *Proceedings of the 7th International Conference on Human Computer Interaction with Mobile Devices & Services*, pages 231–234, 2005.

[dST02]   Vin de Silva and Joshua B. Tenenbaum. Global versus local methods in nonlinear dimensionality reduction. In *Proc. Neural Information Processing Systems, NIPS*, pages 705–712, 2002.

[dST04]   Vin de Silva and Joshua B. Tenenbaum. Sparse multidimensional scaling using landmark points. Technical report, 2004.

[EP05]   Nathan Eagle and Alex Pentland. Social Serendipity: Mobilizing Social Software. *IEEE Pervasive Computing, Special Issue: The Smart Phone*, pages 28–34, 2005.

[FLGC02]   Gary William Flake, Steve Lawrence, C. Lee Giles, and Frans Coetzee. Self-Organization of the Web and Identification of Communities. *IEEE Computer*, 35(3):66–71, 2002.

[Fon97]   Leonard N. Foner. Yenta: A Multi-Agent, Referral-Based Matchmaking System. In *First International Conference on Autonomous Agents*, 1997.

[Gar72]   Eugene Garfield. Citation Analysis as a Tool in Journal Evaluation. *Science*, (178):471–479, 1972.

[GKK$^+$00]   C. Gavoille, M. Katz, N. Katz, C. Paul, and D. Peleg. Approximate Distance Labeling Schemes. Technical Report RR-1250-0, 2000.

[GPPR04]   Cyril Gavoille, David Peleg, Stephane Perennes, and Ran Raz. Distance Labeling in Graphs. *Journal of Algorithms*, 53(1):85–112, 2004.

[GRZL05]   David Gleich, Matt Rasmussen, Leonid Zhukov, and Kevin Lang. The World of Music: SDP layout of high dimensional data, 2005.

[Iwa98]   Yukari Iwatani. Love: Japanese Style. http://www.wired.com/news/culture/1,12899-0.html, 1998.

[Kle98]   Jon M. Kleinberg. Authoritive Sources in a Hyperlinked Environment. In *Proceedings of the Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 668–677, 1998.

[KP05]   Jesper Kjeldskov and Jeni Paay. Just-for-us: A Context-Aware Mobile Information System Facilitating Sociality. In *Proceedings of the 7th International Conference on Human Computer Interaction with Mobile Devices & Services*, pages 23–30, 2005.

[McL64]   Marshall McLuhan. *Understanding Media: the Extensions of Man*. pub-MCGRAW-HILL, 1964. Reprinted by MIT Press, Cambridge, MA, 1994.

[Mil67]   Stanley Milgram. The Small World Problem". *Psychology Today*, pages 60–67, 1967.

[MSN05]   Mehul Motani, Vikram Srinivasan, and Pavan S. Nuggehalli. PeopleNet: Engineering A Wireless Virtual Social Network. In *Proceedings of the 11th annual International Conference on Mobile Computing and Networking, MOBICOM*, pages 243–257, 2005.

[ORT05]   Antti Oulasvirta, Mika Raento, and Sauli Tiitta. ContextContacts: Re-Designing SmartPhone's contact Book to Support Mobile Awareness and Collaboration. In *Proceedings of the 7th International Conference on Human Computer Interaction with Mobile Devices & Services*, pages 167–174, 2005.

[PBMW98]   Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The PageRank Citation Ranking: Bringing Order to the Web. Technical report, Stanford Digital Library Technologies Project, 1998.

[PG04]   Eric Paulos and Elizabeth Goodman. The Familiar Stranger: Anxiety, Comfort, and Play in Public Places. In *Proceedings of the Conference on Human Factors in Computing Systems*, pages 223–230, 2004.

[Pla04]   John C. Platt. Fast Embedding of Sparse Music similarity Graphs. In *Proc. Advances in Neural Information Processing Systems*, volume 16, 2004.

[SM00]   Yasuyuki Sumi and Kenj Mase. Supporting Awareness of Shared Interests and Experiences in Community. In *Proceedings of the International Workshop of Awareness and the WWW*, 2000.

[Sma74]   Henry Small. Co-citation in the Scientific Literature: A New Measure of the Relationship Between Two Documents. *Essays of an Information Scientist*, 2:28–31, 1974.

[TSB98]   Steven R. Thomsen, Joseph D. Straubhaar, and Drew M. Bolyard. Ethnomethodology and the study of online communities: exploring the cyber streets. In *IRISS*, 1998.

[Vos05]   Jakob Voss. Measuring Wikipedia. In *Proceedings 10th International Conference of the International Society for Scientometrics and Informetrics*, 2005.

[Wel00]   Barry Wellman. Changing Connectivity: A Future History of Y2.03K. *Sociological Research Online*, 4(4), 2000.

[Wel05]   Barry Wellman. Community: From Neighborhood to Network. *Communications of the ACM*, 48(10):53–55, 2005.