Recommender Systems with Personality

Amos Azaria and Jason Hong
School of Computer Science, Carnegie Mellon University, Pittsburgh, PA
{azariaa,jasonh}@cs.biu.ac.il

ABSTRACT

We believe that in the future, the most common form of recommender systems will be present in a personal assistant. We claim that such an intelligent agent must be personal, i.e., know its user's preferences and recommend relevant content, a dynamic learner, instructable, supportive and affable. We describe the current state of the art and the challenges which should be addressed in each of these agent properties and provide examples of how we expect future personal agents to convey these properties.

1. INTRODUCTION

Modern smartphones come with an array of rich sensors, fast networking capabilities, significant CPU power, and mass storage. Most smartphones also come with a mobile agent: Android devices come with Google Now, iPhone devices with Siri, and Microsoft devices with Cortana. Despite this, the major use of these intelligent agents is for web search, or for very simple commands such as launching an app or setting an alarm. These agents are not that often used as a recommender systems, and usually do not recommend content without explicitly being asked. Much of the functionality these agents offer does not push far beyond today's direct manipulation user interfaces [8], in which an agent's actions are an immediate response to a user command, or a preprogrammed feature, rather than people using the agent as an actual proxy.

While prototypes for personal assistants date back to over two decades ago [11], we believe that the time is now ripe for making a quantum leap in what an intelligent personal assistant can offer. While, currently, some personal assistants are used for recommendations (e.g. asking Siri for good restaurants near by), and some recommender systems try adding agent personality to them (e.g. Emmy at TasteKid), we believe that future recommender systems will have a much richer "personality" and that the most common form of recommender systems will be a personal assistant. We claim that a personal assistant agent should be: **personal**, being personalized to each user knowing its user's preferences and recommending relevant content, and doing so in a privacy-sensitive manner; **a dynamic learner**, continuously learning and acquiring knowledge about the world and about the user, actively learning

the user's preferences; **supportive**, providing long-term assistance, and initiating actions in order to pursue users goals and tasks; **affable**, offering interaction methods which appear natural to the user and can build rapport with her over time; and **instructable**, allowing the *user* to tailor the agent and its recommendations to her personal needs, by allowing the user to *teach* the agent what she wants and how to do it.

Such an agent is nearly possible due to major advances in hardware, systems, algorithms, and data. Smart devices (such as smartphones, smart-watches, smart-glasses, and smart clothes) are starting to become ubiquitous, and each of these devices comes with a rich range of on-board sensors that can gather tremendous amounts of data about their users and the environments in which they operate. The widespread availability of cloud computing, fast wireless networking, and cheap storage further enhances these capabilities. With respect to algorithms and data, the rise of big data, along with relevant machine learning algorithms, provides us with important capabilities which were previously out of reach, such as high-quality automatic speech recognition, image and gesture recognition, machine translation, and natural language understanding and grounding. There has also been a leap in text-to-speech and non-verbal gesture generation for virtual agents. All these advances provide a never ending interaction opportunity for personal agents to learn user's preferences and provide much better recommendations which are much more personalized to each user and tailored to her needs and goals. In this paper, we drill down on each of the themes we believe an intelligent personal agent should be able to offer within a decade and show how these themes are relevant for recommending better content, actions and behavior.

2. A PERSONAL AGENT

The key idea behind recommender systems is that people are different, and that recommendations should vary according to users' personal preferences and interests. In the past, the content on the web was identical for all users (no personalization), in the present, the recommendation system is identical to everyone (but has personalized output depending on the user's history). We believe that in the future, each user will have her own agent which provides these recommendations. In order for a user to feel that recommendations are truly tailored to her they should come from an agent which is truly personalized to her. Using all the sensors and user history as input, the personal agent may collect enormous information on the user. High personalization level of such personal assistants entails that each user may actually have a unique agent. We therefore propose that each agent may have a unique name, voice and body appearance which travel across all possible platforms (phone, watch, glasses, laptop, etc.). Perhaps, these unique names may be non-word names (which can't collide with other

people the user knows). These names may be generated based on the user's demographics. One name generation method could be based simply on the frequency of different n-grams in names in the user's country, possibly filtered by age and gender[1], or even take into account the user's email, contacts and other personal information. Such agents express the differences between different agents, and make it clear to a user that her agent belongs to her which may help it become more trustworthy.

Since personalization requires collecting sensitive data on the user, it raises a serious threat of privacy invasion. Furthermore, since people are different and unique. Indeed, currently, any usage of an agent comes with a high privacy price tag. Anyone using an agent agrees that any information sent to the agent will serve the company which built it. This implies that people are likely to avoid sharing sensitive information with it. The agent does not actually belong to them, but to some corporation. We believe that this is one of the main roadblocks of this field, and that it will grow stronger as the field continues to advance. We therefore argue that the single overall goal for the agent must be assisting their user and providing relevant content to them in order to help her accomplish her own goals (e.g. saving time, making tasks easier, being more productive, healthier, success in school or other user-defined goals). That is, the agent is there for its user, and data collected by the agent should not be shared with anyone (as if it does not leave the user's phone), unless the user explicitly wants to share this information, or the agent believes that it is the user's will. Using an agent should require a minimal privacy price tag. The current basic market model is that recommender systems are hosted on a website and collect as much information as they can on each user, in order to provide relevant recommendations. Personal agents collect as much information as they can which is shared with the company building those agents, allowing that company to post relevant ads. We believe that if the agent would collect this information and share it will the relevant websites or services on behalf of the user, both sides could benefit. The user will have full control over what is shared with whom, and since the user would receive recommendations which take into account far more information than any singe website could collect, the user is more likely to make a purchase or consume relevant content. While, we may want to assume that all the user specific information does not actually leave the user's phone, this may not be optimal in terms of speed and functionality especially if the agent operates across multiple platforms. Therefore, the user may either have her own hardware which keeps this information, or ,if this information does reside on a companies server, the user must be assured that it is not shared with anyone (not even the agent developers) unless instructed by the user. Another privacy related challenge is selective sharing: since the agent will have access to a large amount of sensitive data, the user cannot always explicitly tell the agent which data to share with whom. Therefore, the agent must learn which information to share with other people, personal agents, or third part agents and apps. The agent is not limited to sharing raw data, but may decide to share (and even sell) some high level, averaged or noised data, share a model built using the data, or provide anonymized data and ensure that the released data may not be de-anonymized. In our vision, there is no single source who has all the data, but many agents each with a different goal, some personal agents and some may be third party agents and even research agents (e.g. a diabetics agent which collects data from diabetics user who decide to share their data).

---

[1]This unique name generation method using letter-based trigram frequency, based on female names, with year of birth of 1980, in the USA, yield the following unique names: Jerlessey, Roberica, Vinaly, Jerin, Pathley, Amarissa, Laniquela, Jillistin and Trany.

Each agent accesses different data, and knowledge is on a "need to know" basis.

3. A DYNAMIC LEARNER

In order to recommend highly relevant content the agent must be a constant learner. The main dimension of learning should be about the individual. Modern smart devices offer a wealth of information about individuals, including one's calendar, Facebook, email, text messages, call logs, and a great number of sensors. This kind of information can be used to learn about a person's communication habits, relationship with others, and general activities, which in-tern could be used to provide more relevant content or recommendations. However, one major research challenge is that it can be hard to create generalizable models of individuals. For example, it is easy for anyone in the world to verify many general inferences (e.g. "Hockey is a sport"), but very few people can verify a personal inference such as "John is the user's best friend". Another major research challenge is that data is siloed across many different devices, apps, services, and formats. As one example, Wiese et al. [15] have found that a person's contacts often spread across their smartphone, Facebook, LinkedIn, and email. Wiese et al. have pointed that it may be hard to integrate contacts across these services, in part due to unusual contact names such as "Pat (Neighbor)," "Mom at Home", and "Do Not Answer". We believe that learning both user preferences and a general knowledge should be an active and never ending process.

Learning should not be limited to the individual user's preferences, but the agent could also better perform if it learned general facts about the world. One example for a never ending learner is NELL [12], which has the goal of reading and learning from the mass amounts of text on the web. The same can be done with the large quantity of images [4], sounds, and videos on the web. This kind of knowledge can help disambiguate queries, understand relationships between entities, common-sense, and help with execution of tasks. For example, if the agent is given a task to organize a surprise party for John, it should know that it should contact John's friends and *not* John.

The widespread adoption of crowdsourcing also offers new opportunities for learning. An agent that requires certain information (perhaps for completing a certain task) might post it on a question answering websites or on a crowdsourcing platform, and have other crowd workers rate the quality of the response. Over time, the agent might also learn from these crowd responses in an active manner. One example for a system that allows people to converse with the crowd is Chorus, [10]. Chorus uses the crowd to rank its own candidate responses in an attempt to provide the best response to a user's query. These responses are later used to train the system.

A simple usage of this knowledge is annotation of content with respect to user goals and preferences. For example, assume that the agent knows that the user is interested in "space exploration" the agent should know that an article about "the curiosity rover" is highly relevant, despite not mentioning the word "space" or even "planet" or "Mars" in it. The agent may use its learned knowledge to encourage the user to perform certain tasks, for example, the agent may tell a user aiming at wight loss, that her friend has just posted that he is looking for a partner for running, or the agent may notice that it is currently raining outside and thus suggest indoor activity instead. When communicating with the user the agent can also refer to common ground, for example, the agent can use a landmark, when assisting the user in a navigation task (e.g. "drive past your grocery store").

Perhaps, the most important knowledge which the agent must pursue to acquire is user preferences. Learning users' preferences

is not limited to user history purchase, or a thumbs up or down on a given content. An intelligent personal agent must proactively try to understand its user's preferences. Therefore, the agent should use active learning and suggests content which may not be the content which the agent predicts that the user will like best, but by suggesting this content the agent learns more about the user, and future content can be better personalized. User preferences can also be learned using preferences of social groups, using common methods such as collaborative filtering [2]. The agent can approach the exploration and exploitation problem in the larger context of all users and try to maximize the social welfare. User preference elicitation is not limited to content provision, but also has to do with learning users privacy preferences. The agent will learn the user's privacy preferences (using similar techniques) and consider these preferences when deciding whether to share some information with a service provider or other agents. Since user models are complex they should not be treated only as a set of weights to be assigned to a set of features. The intelligent agent should try to adopt a set of rules which best explains user actions and the content which the user is most interested in (possibly using first order logic) [14]. The user should be allowed to update this user model and rules according to her actual preferences and interests.

Another challenge for the agent is to learn user routines and usage patterns by observing user behavior. These routines and patterns can be used to discover irregularities, which may impose that the user may require help, in which case, the agent could recommend some actions such as visiting a physician, recommend some different behavior, such a healthier diet, or a relevant product which may be of help. The agent might also recommend a pattern used by a different user, if it assumes that it is more efficient. User usage patterns can also be useful when considering a perfect timing for the agent to interrupt the user for suggesting content, triggering reminders, or obtaining feedback.

4. SUPPORTIVE

We believe that a personal assistant should be a long-term assistant, i.e. take into account the user's long-term goals while providing recommendations or taking actions, just as human assistants may do. A human assistant seeks relevant content and accomplishes tasks in order to complete the requester's long-term goal. Current recommeder systems are very myopic, most of the time recommending a specific product without taking into account that the user may be returning to the website the next day. While some recommender system do try taking the future into account [13], still they are very "narrow minded", limited to recommending a set of products or content, and are totally unaware of the broader context, which includes what the user actually wants and her goals. Current personal assistants, on the other hand, only respond to user queries, actions or specific events. For example, the user may ask a question, and get a response, or send a message, set an alarm or any other action which is immediately executed. The agent should learn about many possible long-term goals in general, but may modify the common time-line structure of a specific goal to best suit its user. A long term assistant must be aware of some of the user's goals and tasks, and initiate actions in order to pursue the accomplishment of these goals over a period of time.

One example of such long-term assistant may include writing a paper for a conference. The agent is aware of the paper content, the co-authors, the deadlines and the venue and monitors the user's progress. The agent can recommend related work and notify the user if new relevant papers are published. It can show a dashboard helping the user track the progress in writing the paper and alert the user when proposed milestones should be reached and make sure that the user notices relevant emails. If required, the agent may recommend that the user contacts co-authors who were asked to provide feedback or write certain parts, if it notices that this hasn't been done. Eventually, the agent may help in registering for the venue and recommend a hotel and flight. The agent can also help with reimbursements before or after the trip. The agent can apply knowledge it has about trips in general for providing recommendations (such as interesting places to visit, restaurants, hotels etc.). Additional long-term examples may include, losing weight, health behaviors such as workout and taking medication, teaching a class. A long-term goal may even be buying an item, such as a good camera, in which the agent may recommend content which will allow the user to learn about the possible options and important properties of the interested item (e.g. ISO, HDR, Shutter-speed, focal length, aperture etc.), and only then recommend content related to specific items or brands. With such goals which span over months or even years, it may be challenging for the agent to acknowledge whether it was helpful, and whether its provided content was useful. Depending on its belief, it may decide to selectively share some information with other users' agents.

5. AFFABLE

Human communication and rapport establishment involves many channels, including verbal speech, tone, and body language. The timing of speech also plays a role, with people cutting-off or completing each other's sentences. Humans also take into account each other's context and will talk differently if they know, for example, that someone is currently driving. However, today's agents only make use of very little of this rich input. Recent advances in gesture detection, acoustic modeling, and virtual agents make it possible to offer an embodied agent which can take in as input both audio and visual non-verbal cues [5]. The agent can also incorporate multiple approaches to establish rapport with the user, for example by varying its language and gestures as a function of the rapport level it believes that it is having with the user [6]. This same agent should be ubiquitous, and allowed to be accessed via wearable devices, laptops, tablets, etc. Cross-device presence raises a major challenge for a consistent and unified user experience. The agent must be fully aware of its actions and which interaction channels are currently being used. For example, if the user is using a navigation application, and tells the agent "I can take it from here", the agent should reply gracefully and stop providing navigation instructions (unlike current agents which are not aware that they are currently providing navigation instructions, and thus perform a web search instead). The agent also needs to appear as unified to users, as if it were a single entity, despite having a wide range of front-ends that users interact with. The agent also needs to capture what interactions have happened on different devices, and also be aware that it may not be able to perform certain tasks on certain devices.

6. INSTRUCTABLE

Current personal agents are rigid, limited to predefined commands, and not extendable by the end users. For example, if Google Now finds a flight reservation in a user's email it will recommend the user to leave for the airport when it is time to do so. However, this capability is predefined (handcrafted) and the user cannot extend this functionality by requesting the agent to notify her if she should leave for an important meeting located out of town. Current apps are developed only if there are enough users who will use them.

While this theme falls beyond recommender systems, our vision is to allow end users to instruct their agent in natural language, us-

ing the agent's sensors and effectors as well as its general and private knowledge. The user should be able to define any procedure or rule (if-then clause) in a natural way, just as the user would instruct and teach another human [1, 7]. Recent advances in grounded language acquisition [3], learning by demonstration [9] and natural language parsing and understanding, lead us to believe that overcoming this challenge is closer than ever.

If we take the above example, a user should be able to define her own command simply by saying: "let me know when it's time to leave for the airport". The agent may know how to add this rule, but if not, the agent may ask follow-up questions leading the user to further explain. Users will be able to define numerous functionalities without a need for knowledge on programming. Once functionalities are defined, users will be able to share these functionalities with each other.

We argue that the agent should learn these commands on demand and in a natural way for the user, i.e., the user gives commands (or creates rules) and if the agent doesn't understand, it may ask the user to explain what it should do (or how it should apply the rule). With such methodology, the user is not required to inform the agent that it is about to teach it something new, but simply uses the agent casually, and only if the agent fails, will the user be required to teach it. Since the user actually needs this functionality (and thus gave the agent the command), she may be willing to spend the time and teach the agent this new functionality. Once acquired, the user may share this new acquired command with others or the public. If, on the other hand, the user refuses to teach the agent or fails doing so, the agent may contact the crowd for help. Some users (or crowd workers) may also help by defining composite virtual sensors which can be used by others.

7. CONCLUSION

In this paper we present our vision for a personal assistant, argue that in the future, the most common form of recommender system would be a personal assistant, and list the most important properties of such an agent. The question remains, as whether a personal assistant can be viewed as a form of a recommender system. We believe that any boundaries between the two, lie in the level of autonomy given to the agent.

In the present, the most commonly used recommender system, is one in which the recommender system responds either to an explicit or implicit request by a user, and provides either recommended items, content, movies, etc. The next level of autonomy, could include agents that initiate *action* recommendation such as recommending that the user leaves to the airport (as present in Google Now). The next level can include an agent that takes actions on behalf of the user but still requires confirmation on any such action (for example, any email it sends on behalf of the user). All actions of such an agent, could still be viewed as a recommendations. The next level of autonomy, includes an agent which may autonomously take actions on behalf of the user if it is confident enough that the user would confirm the action. Future personal assistants may either have the user set a threshold for this confident level, or learn the optimal threshold for each user and each domain (based on user actions and data). We believe that this intermediate-level of autonomy is the perfect blend of recommender systems and fully autonomous agents. Indeed one of the original papers to describe the use of personal assistants, took use of such a threshold and performed actions autonomously, if such a threshold was met [11]. A fully autonomous agent which acts on behalf of the user at all times (possibly leaving the user some high-level control), is no longer a recommender system. However, we believe that while such level of autonomy may be useful for some domains (e.g. space exploration), it is not very useful as a personal assistant, as we believe that even in the far fetched future, users would always want some level of direct control over their personal assistants.

9. REFERENCES

[1] Amos Azaria, Jayant Krishnamurthy, and Tom M Mitchell. Instructable intelligent personal agent. In *AAAI*, volume 4, 2016.

[2] John S Breese, David Heckerman, and Carl Kadie. Empirical analysis of predictive algorithms for collaborative filtering. In *Uncertainty in artificial intelligence*, pages 43–52. Morgan Kaufmann Publishers Inc., 1998.

[3] David L Chen and Raymond J Mooney. Learning to interpret natural language navigation instructions from observations. *San Francisco, CA*, pages 859–865, 2011.

[4] Xinlei Chen, Ashish Shrivastava, and Arpan Gupta. Neil: Extracting visual knowledge from web data. In *Computer Vision (ICCV)*, pages 1409–1416. IEEE, 2013.

[5] Fiorella De Rosis, Catherine Pelachaud, Isabella Poggi, Valeria Carofiglio, and Berardina De Carolis. From greta's mind to her face: modelling the dynamics of affective states in a conversational embodied agent. *International journal of human-computer studies*, 59(1):81–118, 2003.

[6] Lixing Huang, Louis-Philippe Morency, and Jonathan Gratch. Virtual rapport 2.0. In *Intelligent virtual agents*, pages 68–79. Springer, 2011.

[7] Ting-Hao Kenneth Huang, Amos Azaria, and Jeffrey P Bigham. Instructablecrowd: Creating if-then rules via conversations with the crowd. In *CHI*, pages 1555–1562. ACM, 2016.

[8] Edwin L Hutchins, James D Hollan, and Donald A Norman. Direct manipulation interfaces. *Human–Computer Interaction*, 1(4):311–338, 1985.

[9] George Konidaris, Scott Kuindersma, Roderic Grupen, and Andrew Barto. Robot learning from demonstration by constructing skill trees. *The International Journal of Robotics Research*, 2011.

[10] Walter S Lasecki, Rachel Wesley, Jeffrey Nichols, Anand Kulkarni, James F Allen, and Jeffrey P Bigham. Chorus: a crowd-powered conversational assistant. In *UIST*.

[11] Pattie Maes et al. Agents that reduce work and information overload. *Communications of the ACM*, 37(7):30–40, 1994.

[12] Tom M Mitchell, William Cohen, Estevam Hruschka, Partha Talukdar, Justin Betteridge, Andrew Carlson, Dalvi Mishra, et al. Never-ending learning. In *AAAI'15*, 2015.

[13] Guy Shani, Ronen I Brafman, and David Heckerman. An mdp-based recommender system. In *Uncertainty in artificial intelligence*, pages 453–460. Morgan Kaufmann Publishers Inc., 2002.

[14] William Yang Wang, Kathryn Mazaitis, and William W Cohen. Programming with personalized pagerank: a locally groundable first-order probabilistic logic. In *Information & Knowledge Management*, pages 2129–2138. ACM, 2013.

[15] Jason Wiese, Jason I Hong, and John Zimmerman. Challenges and opportunities in data mining contact lists for inferring relationships. In *Proceedings of Pervasive and Ubiquitous Computing*, pages 643–647. ACM, 2014.