Defining Groups:  
Identifying Characteristics of the Mars Scientific Community

Pamela R. Dowling  
Wright State University

Charles J. Budney  
Jet Propulsion Laboratory,  
California Institute of Technology

Deborah S. Bass  
Jet Propulsion Laboratory,  
California Institute of Technology

Abstract

Scientific communities are composed of interconnected and disparate groupings of individuals. The connections between those communities and the backgrounds of individuals can influence how those groups approach and solve problems. A question then becomes, what aspects of their connections and backgrounds are important to characterize that community? How do you characterize a community that spans multiple disciplines, locations, and research interests? The current research sought to identify defining aspects of the Mars scientific community by looking at connections represented by co-authorship on papers. A subset of the Mars community was studied, specifically, those taking part in two separate panels. These panels consisted of the recent (2009) and previous (2003) Planetary Science Decadal Survey for Mars. The data suggests the community holds a diverse background with a fair amount of crossover in their academic achievements. Results also point to the influence of universities in the composition of these two panels.

1. Introduction

The idea of exploring the planet Mars has attracted the attention of many members of the scientific and engineering communities. Exploration of our nearest neighbor planet has proven to be a highly involved venture including large numbers of people with differing backgrounds and research interests. Those varied areas of expertise work together toward the common goal of pursuing science and engineering in a new frontier of space exploration.

It is important and beneficial both to those who are members of such groups as well as members of oversight to know the characteristics of communities to better understand them. To characterize the make-up of a community, it is important to look at the people who make up that community. Individuals bring with them differing areas of expertise, backgrounds, and perspectives which all add to the community [1, 14].

The goal of the present research supported by the Mars Exploration Program Science Office was to take a closer look at the composition of scientists who are interested in Mars science. The general method used to accomplish this task was a social networking analysis [7, 9, 10], though the people included in this analysis may not constitute a readily identifiable and consistent community. Instead, the attempt here is to identify a method for recognizing a community that has been formed even if its members have minimal if any contact with each other.

The project undertaken here aimed to collect information that would give a picture of those interested in the broad area of Mars science. The goal was to represent those who share intellectual interest in Mars in a way that would give us a better understanding of the dynamics and structure of this group of individuals. Before taking this step, the question of how one characterizes any community needed to be addressed. How do you represent a scientific community? What aspects define people who gather for a common goal?

There are many possible ways to characterize people in a community. In general, one approach to understanding a group is to look at characteristics of individuals in the group. Age, race, and gender are typical ways to define individuals in a group or team [6, 13] and are often used in identifying group characteristics. These demographic categories have a tendency to be superficial so other defining characteristics were pursued in the current research. One might define a group with a map of an individual’s geographical location. This information may show cultural perspectives and governmental influences that could influence dynamics in a group. Location might also impact communication patterns within a group [15]. Another area that could be used to describe a group is research interests. Mars scientists all share an interest in the planet Mars, but interests are shaped by research area. Whether someone is a geologist, an astrochemist, or an astrobiologist will affect how they approach...
situations. Experience on past projects and past places of employment impact what people bring to a group. Another individual influence is academic heritage, including schools graduated from and advisors worked under. Each of these areas represents potential ways to characterize a community.

In social network analysis, research points to communities of practice as organizing entities where one can begin to identify group membership [5]. These communities are locations where learning and apprenticeship occur [12]. Though traditionally this would refer to a group that had several shared features, such as commonly, a shared geographical location or method for accomplishing a goal, it might be stretched to describe a more distally organized group as we have in those who study Mars.

Social network analysis also addresses the strength of ties between individuals as a way to characterize a community. When people regard others as integral to their social, cultural, or work relations they are thought of as strongly connected. Those who are considered less relevant to daily life are more weakly tied. One can also talk about people who are central to a group and those who retain a marginalized position [9].

Ties between individuals can also be understood by looking at how members of a community are connected to each other. Connectedness can be a difficult concept to measure as many social interactions are informal and hard to track. One way of defining connections between people is to look at those who have written papers together. The scientific community traditionally communicates their thoughts and accomplishments through peer reviewed journal articles that often are authored by more than one individual. Looking at how members connect by authorship can give insight into group affiliations and overlapping research interests. It can show professional relationships between individual community members and even include relationships that might not have been remembered if the people involved were asked directly [9]. We can see the breadth of the community as well as the level of interconnectedness by looking at this method of characterizing a community.

To analyze all people who have scientific or engineering interest in Mars is a daunting task. One reason is they do not readily lend themselves to identification as a community. Those interested in Mars are affiliated with multiple organizations, scientific disciplines, and country citizenships. Also, Mars missions span decades of time and include thousands of individuals who have worked or collaborated on more than 20 different missions [11]. (This number represents the missions undertaken by the efforts of the United States. When including other countries, the number increases to over 40 missions). Each individual who has participated in Mars exploration could have been a member of any number of committees, groups, or projects.

Rather than take on the whole of those interested in Mars exploration, a smaller subset of the community was the focus. The data was first narrowed by looking only at those interested in Mars science, as opposed to those who would fit into the broad categories of computer scientists or engineers. Even with this criterion, the number of individuals composes a very large group.

One set of individuals who offered an opportunity to narrow the field of view were members of an influential group known as the Mars Panel of the Planetary Science Decadal Survey. The subset of individuals in this group has essentially been selected as representatives of all of the scientist and engineers interested in Mars. This group is formed regularly but each specific panel only meets a few or more times throughout a one year time span.

There is evidence to suggest that individuals in a loosely connected group such as those interested in Mars science and more specifically, those partaking in the Decadal Survey for Mars do indeed influence each other, even if they do not have direct contact with each other [8]. By mapping out a portion of this community and showing the connections made through co-authorship, we can see some of the pathways that information may flow through [4].

The most recent two panels, 2003 and 2009 were used here. The members of each year’s panel have been put in a position of power where they could set policy for future Mars science and exploration. For instance, they encouraged research into the geological history of the planet Mars. This group, already selected as representatives of Mars science, was identified as a community of interest warranting more in depth analysis.

2. Method

The data collection for the current research consisted of identifying and collecting information related to publications and authors of those publications. Data collection was focused on authorship lists for those who have taken part in the Mars Panel of the Planetary Science Decadal Survey. Information on co-authorship came mostly from online CVs and personal websites. The term co-author is used to represent any name on a paper that included a Survey member as an author.
Online CVs were used as a primary source of information for finding co-authors of peer-reviewed journals. When publication lists were not found from online CVs, the search engines Science Citation Index or biomedexperts.com were used to find articles written by select individuals.

Any additional information pooled from the subjects in the search came mainly from their CVs as well as web searches (e.g. google searches) and personal web pages. NASA project webpages were another source of information. These project pages post short bios on the people involved in various missions. Another website that was used to collect data is linkedin, a social network site for professionals where basic school and work histories are voluntarily provided by users. Although this information was collected, the main data presented here include the information found on co-authorships of papers.

Author lists were pulled from online CVs and other web based pages without regard to an authors’ ranking on a paper. Whether the author was first, second, or last author, all of the authors listed on papers that included a Survey member were included in the analysis. The authors on all papers found were included so that the length of a person’s research life was represented in the data.

Once the data was collected, excel worksheets were used to organize, characterize, and validate the information. In arranging the lists of those people who co-authored papers with the target people, duplicate entries were deleted such that if the same name appeared more than once as a co-author on a single Survey member’s list, their name was only counted once. Some attempts were made to adjust for typos in the original sources. For instance, if the name ‘S. F. Lastname’ was listed multiple times and the name ‘S. G. Lastname’ appeared only once, a google scholar search of ‘S. G. Lastname’ was conducted. If the person did not appear to exist or have even remote relevance to the space community, the middle letter was assumed to be a typo. Names of co-authors were matched with Survey members and aggregated. Initials were later removed from the lists unless they were necessary to differentiate between different authors. This step was performed to simplify visual representations of the data.

Social network analysis can involve a very complex set of formulas [16]; however, the current analysis used a simplified approach to social network analysis and tools that visually represented the connections between data points.

After lists of information were collected, they were entered into social network software programs called UCInet and NetDraw [2, 3] which allowed for visual representations of data. Planetary Science Decadal Survey members were designated as nodes, and co-authors were linked to these nodes. The visual representation of nodes and links connecting the nodes were shifted until patterns became clear.

3. Results and discussion

There were 26 members of the 2003 and 2009 Mars Panel of the Planetary Science Decadal Survey (14 for 2003 and 12 for 2009). The 2003 list included the chair and vice chair of the Planetary Science Decadal Survey as well as the 12 members of the Mars Panel. The publication lists for 18 of the Survey members came from online CVs (see Table 1). Co-author data for six Survey members came from online search engines (Science Citation Index or biomedexperts.com). For two of the members of the 2003 committee, no lists of publications could be found from online CVs. Additionally, as these two names were fairly common, pulling information from a source like Science Citation Index was not a reasonable option. These two people are not represented in the data. With the loss of these two names and the addition of the chair and vice chair of the Survey, each Planetary Science Decadal Survey panel studied here had 12 members.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Information sources</th>
<th>2003</th>
<th>2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>Online CVs</td>
<td></td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>Science Citation Index</td>
<td></td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Biomedexperts.com</td>
<td></td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>No data</td>
<td></td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

The total number of unique co-authors represented in the data comes to 2733 (1082 people for 2009; 1845 for 2003 – these numbers do not add to 2733 because of crossover between the years). The minimum number of co-authors a Survey member had was 12. The maximum number of co-authors was 368. The average number of co-authors per Survey member was 114.

In Figure 1 (p. 8), members of the 2003 Survey are represented in green, 2009 Survey members are represented in red, and co-authors are shown in blue. Some of the links between data points show that people in the 2009 Survey tended to share co-authors with other 2009 Survey members. There is also a tendency for 2003 Survey members to share co-
authors with other 2003 Survey members. However, we also see cross year connections. Many co-authors bridge the distance between the 2003 and 2009 Survey members.

3.1. Single connection nodes

Some of the co-authors represented in the data arrange themselves into large clusters around the Mars Panel of the Planetary Science Decadal Survey members. These clusters show many co-authors who connect to one Survey member, but not to any of the other Survey members. These single connection nodes show that Survey members have connections with individuals outside of their peers among the Survey. This data gives a picture of the breadth of the community represented by the Survey members in that these links suggest the individuals on the Survey are researching areas that are apart from their Mars research. Though it is possible that this represents a lack of scientific communication between members, it is highly likely that these members simply research in areas that go beyond the interest of other members of the Mars Panel of the Planetary Science Decadal Survey.

From a visual analysis, there appear to be more of these large single connection clusters around those from the 2003 Survey (green nodes) than the 2009 panel (red nodes). This suggests that perhaps there are slightly more members on the 2003 panel whose members are touching on research areas that span beyond the interest areas they would share with other Survey members.

This breadth of research can have both benefits and drawbacks when it comes to the research community focused on Mars. These who are less connected with highly involved individuals might be described as marginally connected to the Mars scientific community [9]. When someone is marginally connected, there are implications for how willing that person is to adopt new methods (generally, they are more willing) as well as influence how likely those members are to share information, and how they interact with others in the community [5, 9]. When members partake in a more centralized position, as will be described next, they also are likely to act in ways that are a reflection of their standing within the community.

3.2. Multiple connection nodes

The data also shows a large number of links connecting co-authors to many Planetary Science Decadal Survey members. This can be seen in the blue clusters scattered throughout the picture. In these connections, we can see patterns of data where groups of co-authors connect two or more Survey members.

The red nodes appear to have more nodes that connect to more than one Survey member. Though they certainly have groups of people they co-author with who are not connected to other panel members, they appear to represent a somewhat higher level of connectedness than those who took part in the 2003 Survey.

Among those co-authors who connect to more than one Planetary Science Decadal Survey member, we can see many co-authors who have connections with 3 or more Survey members. These have been arranged into the star patterns toward the center of the picture. These co-authors appear to be well connected to other Survey members. When we see Survey members linked to many co-authors who are also linked to other Survey members, we can see patterns in interconnectedness.

The data suggests that people who are a part of the Mars Panel of the Planetary Science Decadal Survey are linked to many people through co-authorship. The large number of associations suggests the span of their scientific connections. We also see that Survey members wrote papers with many of the same individuals that other Survey members wrote papers with. This finding suggests that there is a high level of overlap and interconnection between Survey members. Rather than an individual scientist who forges on in his or her own area of interest, scientists here are communicating with and partnering with other scientists.

Survey members might be considered centralized figures in the Mars scientific community along with those that are highly connected to Survey members [9]. When looking for someone to represent scientific interest in Mars, it would be good to include those names of people who are highly connected. This will ensure those selected people are aware of current research as they have been personally involved in much of it.

When comparing the potential implications of multiple connection nodes with single connection nodes, one might think about the burden and benefit of spreading yourself over multiple areas of research. Those who are more connected with other Mars scientists in this data likely focus their research more narrowly. When research is focused on a small subset of information, you tend to be quite competent in that area, but may lack in areas outside. When you cover multiple areas of study, as suggested by those people who had many single connections with nodes.
that do not connect with others, you may fit yourself into the colloquium of the ‘jack of all trades, master of none.’

3.2.1 Influence of agency. Additional analyses were performed on the data that had been collected. In Figure 2 (p. 9) data was pooled for those people who had multiple (three or more) connections to different co-authors. The people in this analysis are those who form the blue starbursts near the center of the figure where the nodes connect to multiple decadal Survey members.

In this analysis, the organizations where the Planetary Science Decadal Survey members worked were identified as a way to organize the data. The represented organizations fell into 5 categories - Universities, Companies based in the United States, European University, NASA research centers excluding JPL, and JPL as its own category. The connections between the co-authors (who connected to multiple other co-authors) and the organizations were mapped.

Each link represents a unique affiliation with a type of organization or agency. Most of the agencies represented by the Planetary Science Decadal Survey panel members were listed as universities (Table 2).

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Agency affiliation of Mars Panel of the Planetary Science Decadal Survey members</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2003</td>
</tr>
<tr>
<td>US University</td>
<td>7</td>
</tr>
<tr>
<td>NASA</td>
<td>0</td>
</tr>
<tr>
<td>JPL</td>
<td>1</td>
</tr>
<tr>
<td>US Company</td>
<td>4</td>
</tr>
<tr>
<td>European University</td>
<td>0</td>
</tr>
</tbody>
</table>

The data in the table show that many members of the Planetary Science Decadal Survey panels affiliated themselves with universities. It is then somewhat unsurprising that there are many connections between universities and co-authors in the figure (Figure 2). However, it is interesting to note the prominence of universities in this group.

In this data, the affiliations were collapsed so that if a co-author wrote papers with multiple people from universities, those universities were represented as a single link. Even with this condensing of the data, universities play a major role in the authorship of papers that were linked to the Mars Panel of the Planetary Science Decadal Survey members. We even see universities as the only agency (outside of a single node connecting to a US based company) where co-authors link to only one type of agency. Universities represent a major influence on those who are members of the Mars Panel of the Planetary Science Decadal Survey and those who have multiple connections to different Survey members.

3.3. Future work

The work presented here represents a small but vital group of people within the Mars scientific community. To gain a larger perspective on the Mars community, it would be interesting to include more than those who participated in the Mars Panel of the Planetary Science Decadal Survey. Another way to improve on this work would be to explore other ways of characterizing the community. There are many ways to characterize a group and several were mentioned in this paper including past projects, geographical location, and academic background. The present paper narrows in on only one of many avenues available.

Authorships in this research were collected without regard to first, second, or other authorship. It is possible that many of the papers collected here were written by students whose advisor was a member of the Survey panel and who therefore listed their advisor as a co-author. This could be the reason for the large clusters of people surrounding individuals on the Survey panel who are not connected with other members of the panel. Teasing that information apart would be another next step in this research.

The information on co-authors was pulled from peer reviewed articles. This was done as these articles represent a large portion of the written work of most scientists. It also offered the ability to focus the scope of the analysis even by a small amount. It would be interesting to have included conference publications and other professional activities of those involved.

Because the data collected here was pulled from authorship lists, the only people representing the Mars community for this project were those who wrote paper(s) with Survey members. This does not give an expansive view of all those involved in Mars research and this limitation produces its own form of confounds. Very likely people who write papers together also have other factors in common such as physical proximity through work affiliations, affiliation with the same university, or being a student under a shared advisor. In addition, authors who were not mentioned on CVs were not included,
therefore there is nothing to compare these findings against if an attempt were made to link these names to all those interested in Mars. A larger scale study would be able to more adequately address the desire to characterize the whole of the Mars scientific community.

It would also be interesting to see how many of the people represented in the co-authorship lists were people who have been identified as members of the Mars scientific community or if these people represent interest areas outside of Mars research.

If this research were to include more years of Mars Panel of the Planetary Science Decadal Survey members, both past and future, we may see more trends emerging. Perhaps the number of connections between individuals has been impacted by technology which has allowed greater communication between individuals.

There is clearly more research to be done to further the growing picture of the Mars community. It is interesting to note that the analysis of co-authorship on papers do give some helpful information about the interconnections between group members. This method may be used to describe groups in other arenas where the community presents its work through peer reviewed journal articles.

4. Conclusion

The present research focused on a method for indentifying the existence and boundaries of a community even when that community covers individuals from vastly different backgrounds and perspectives. By looking at co-authorship on papers from a subset of Mars scientists, a characterization of the community was possible. Though there are many ways to depict a group, focusing on authorship in journal articles can show relationships between people and give a picture of the underlying structure of a community.

We can see from the sheer number of people represented in the co-authorship data that the Mars Panel of the Planetary Science Decadal Survey members are connected to a wide net of people. This piece of information gives support to their selection as members of this panel.

The connections linking single co-authors to multiple Survey members suggest the Survey members are well connected to other members in their community. Those on the Mars Panel of the Planetary Science Decadal Survey were selected to represent those interested in Mars science. The results of this analysis suggest that they do a good job of representing their community.

Analysis of the data also shows a large connection to people affiliated with universities. The influence of universities should not be ignored in the group analyzed here. Many of the members of the Mars Panel of the Planetary Science Decadal Survey present themselves as members of a university. Though they clearly affiliate themselves with the scientific study of Mars, their primary connection is with a university. The prominence of universities in the Survey additionally shows up in the analysis of those co-authors who wrote papers with three or more Survey panel members. Even after collapsing the university affiliations so that it is only represented once per co-author, universities still show a large number of connections.

The data presented here presents research that was gathered with the aim of creating a representation of the Mars scientific community. The data on even a small subset of that community (namely the Mars Panel of the Planetary Science Decadal Survey members) helped create a picture of the interconnections we see between people which can then help us to understand the composition of this group whose members have been chosen to represent the Mars scientific community.

5. References


Red dots represent members of the 2009 Mars Panel of the Planetary Science Decadal Survey. Green dots represent members of the 2003 Mars Panel of the Planetary Science Decadal Survey. Blue dots represent co-authors who wrote papers with one or more persons on either Mars Panel of the Planetary Science Decadal Survey. The links between the nodes show where people wrote papers together. Created using NetDraw and UCINet [2, 3].
Figure 2. Planetary Science Decadal Survey 2003 and 2009 Mars panel members with 3 or more connections between co-authors, grouped by agency.

Red dots represent co-authors who wrote papers with three or more different members of the Mars Panel for the Planetary Science Decadal Survey members. Blue dots represent the collapsed agency affiliations of members for the Mars Panel of the Planetary Science Decadal Survey. Created using NetDraw and UCInet [2, 3].