

A Conceptual and Operational Definition of 'Social Role' in Online Community

Eric Gleave
University of Washington
egleave@u.washington.edu

Howard T. Welser
Ohio University
welsers@ohio.edu

Thomas M. Lento
Facebook
lento@facebook.com

Marc A. Smith
Microsoft Research
masmith@microsoft.com

Abstract

Both online and off, people frequently perform particular social roles. These roles organize behavior and give structure to positions in local networks. As more of social life becomes embedded in online systems, the concept of social role becomes increasingly valuable as a tool for simplifying patterns of action, recognizing distinct user types, and cultivating and managing communities. This paper standardizes the usage of the term 'social role' in online community as a combination of social psychological, social structural, and behavioral attributes. Beyond the conceptual definition, we describe measurement and analysis strategies for identifying social roles in online community. We demonstrate this process in two domains, Usenet and Wikipedia, identifying key social roles in each domain. We conclude with directions for future research, with a particular focus on the analysis of communities as role ecologies.

1. INTRODUCTION

Social life has moved online. From discussion boards, to wikis, to social networking sites, people do things together through digital communication. Those interactions leave behind complex records of who did what, when, under what context, and with whom. In other words, the interaction order [16] is now electrified and self-documenting. This fact is reverberating across the social, communication, information, and computer sciences and requires the development of new methods and increased interdisciplinary collaboration [39]. This collaboration requires researchers to use the comparative advantages of many disciplines in order to perceive, conceptualize, measure and understand what has unfolded in the electronic interaction order [51]. A key intersection of these scholarly disciplines is the concept of the social role [40] [8] [52] [50].

Whenever people interact, they act in ways that are both enabled and constrained by social structure:

the social context, history of actions, structures of interaction, and the attributes people bring to the interaction. Social roles encapsulate these dimensions by providing a meaning system that defines alternative actions and imposes structure on who chooses to interact with whom, under specific conditions and circumstances [23] [29] [34]. Social roles are a crucial window into the underlying social structure of interaction, and one which researchers can use to differentiate research subjects and explain how and why people behave in the ways they do [8] [11] [17] [18] [19] [31].

We articulate a conceptual definition of 'social role' based on the symbolic interactionist approach to identity theory [43] [8]. That definition describes social roles as cultural objects that are "recognized, accepted, and used to accomplish pragmatic interaction goals in a community" ([8]: 232). Such roles operate as both sources of constraint and resources for action and thus are observable at several levels of analysis: in the content of communication, in relation to the identities enacted, in distinctive patterns in local social networks, and in the behavioral history of community members. Operationally, we suggest that once such roles have been identified through a combination of detailed analysis of content and broader statistical analysis of structure and content, researchers can use patterns at the levels of social networks and behavior to infer role status of participants in online social spaces.

Building on the conceptual and operational definition, we describe measurement and analysis strategies for identifying social roles in computer-mediated social spaces. We then demonstrate how the combined conceptual and operational definition allows the study of communities. We demonstrate this process in two domains, Usenet and Wikipedia, identifying key social roles in each. Finally, we conclude with directions for future research, including analyzing the interactions between various roles in the context of a larger role ecology.

2. DEFINITION OF SOCIAL ROLE

The concept of social role has been the subject of analysis for over 100 years (see e.g. [4] [16] [28] [29] [30] [33] [35] [36] [40]). The term's meaning is therefore broad and varies from area to area. Despite this variation, the concept remains important because of its utility: the classification of types of social relations and behaviors into a smaller set of roles reduces the analytic complexity of social systems and facilitates the comparative study of populations across time and setting [22]. This categorization of users can be particularly useful when designing or managing social media spaces. By identifying roles and mapping participants to one or more role categories, designers and managers of computer-mediated social spaces may focus on meta-level management tasks like monitoring the relative proportions of roles. The creation and maintenance of these social spaces depends on the complex social ecology created by the interaction of several social roles. Identifying roles is a critical step in the effective management of these social spaces.

2.1 Social roles are behavioral and relational patterns

Social scientists draw attention to two fundamental dimensions of social life: structure and culture. Social structure refers to patterns in the distribution of relationships and resources in a population. This includes things like the pattern of friendship relationships in an online community, or the age distribution within a town. Such structural features act as both opportunities and constraints for social action.

Our definition asserts that social roles begin from a structural foundation in simple commonalities in behavior. The role of father begins from actions related to socializing, protecting, and caring for offspring. The particular actions may vary across cultures, but the role of father fills certain basic social needs and is therefore recognizable across cultures, both by outside observers and individuals living within a given social environment. Similarly, many social roles, especially those that are newly emerging, will have distinctive social structural foundations even if they have not yet developed the same level of recognition within and across cultural settings.

An understanding of the structural foundation of roles goes beyond analysis of behavioral patterns. Because behavioral consistencies also pattern relations, people who consistently adopt particular roles develop distinctive modes of participating in social settings, which results in patterns of behavior and relations. Once the correlation between these patterns and a particular role is established through statistical analysis, those patterns can be used to infer role use from those structural signatures [50].

It is important to note that cultural elements are a key part of many conceptualizations of social roles. Callero [8] emphasizes the importance of cultural dimensions, especially for well-established roles that are recognized in social settings. Callero also emphasizes that roles vary in terms of social accessibility (barriers to adoption of the role) and situational contingency (degree that the role requires specific contexts to be used). In other words, roles may differ in their scope and their relevance to action.

Roles are resources that help people accomplish their goals and are tools used in the establishment of social structure [2]. For example, a father in the company of a large number of young children may adopt the role of 'coach' in order to achieve social order and to justify the sanctioning of children who are being disruptive. Similarly, in an online community a participant may rely on the role of 'expert' in order to achieve social order and to sanction 'newbies' for failures to follow local norms (like reading the FAQ). Both of these cases highlight the importance of scope; the role of 'coach' or 'expert' may take different forms in different contexts, even as those roles have certain similar characteristics. More importantly, in many contexts these roles may not be explicitly named. Experts in online discussion groups can fill that role without formal recognition from other members of the group, or even recognizing the roles themselves. We emphasize the importance of finding ways to measure and identify such emergent roles, especially in those cases where the cultural dimensions of a role have not yet been recognized by all of the participants in a given social setting.

2.2 Conceptual operationalization of roles

Social roles can be conceptualized at several different levels of abstraction. At the most granular level, social roles can be defined in terms of behavioral regularities and network attributes. Consistent behaviors resulting in persistent or recurrent interactions between individuals in a social group are potential signals of a meaningful social role. Social actions, loosely defined as individual behavior that includes an action targeted at one or more of the other members of a social group, are a higher level measure which is often useful for identifying the existence of a role. Beyond behavioral metrics, researchers can look for self-identification with a given role.

Once the data from these levels of analysis has been processed, role researchers might come up with a definition for a particular role. That might be a definition for a role which is emergent and as yet unknown to the participants in the social group or one for a fully recognized social position. The researcher

can then move from that role definition to more abstract categories or classifications that provide theoretically or practically relevant information about the purpose and function of a given role.

Table 1 uses the example of a vandal fighter, an important social role within the Wikipedia editor community, to illustrate these different levels of analysis. This role was initially identified through exploratory qualitative analysis of interactions across the dozens of community and project related talk pages in Wikipedia. It was later verified both through more detailed quantitative analysis and through discovery of patterns of self-identification and cultural significance.

We contend that the best way to identify roles is to begin with either qualitative or quantitative analysis of social action and work both downward towards identifying the key related behavioral regularities and distinctive positions (signatures of social roles), and upwards to abstract theoretical categories that allow us to tie these particular roles to general research objectives.

The initial stages of social role research involve an iterative process of moving from observations of meaningful situated actions of actors likely to be using a particular role, to measures of structural and behavioral patterns. There are two related goals: the first is to refine understanding of the role itself and the second is to generate insight into the structural

signatures of that role. Once the behavioral and structural signals of that role are understood, researchers can attempt to discover indicators of social roles at other levels of analysis. As the researcher comes to understand the relationship between those indicators, particularly indicators related across levels of analysis, observations at one level can be used to infer the degree of role use by community members. At this stage, the relative frequencies of different role types can be used to characterize the community more generally (for a more detailed discussion, see section 5 on role ecologies).

It is important to choose the right starting point for social role research. Starting from simple behavioral regularities or distinctive social positions is a flawed approach because it is overly inductive: of the infinite patterns we can find, there is no reason to think that the ones that initially stand out will be of any social significance. Starting from abstract categories like 'altruist' commits the error of over-deduction: simply because we can label a set of activities as altruistic does not necessarily mean that those behaviors are motivated by altruism. The key is to connect higher levels and abstract categories to observable behavioral regularities and distinctive network positions in order to discover types of actors relevant to theoretical questions in social research.

Table 1. Operationalizing roles across levels of analysis

| Level of Analysis | Operationalization | Example: Vandal Fighter in Wikipedia |
|--------------------------------------|--|---|
| Behavioral regularities | Persistent, repeated actions | Vandal fighters exhibit numerous reversions of vandalized pages, a serial pattern of edits to the user pages of other users, and edits to vandalism related project pages |
| Network attributes | Distinctive, reproducible patterns of interaction across several examples of the role type | Vandal fighters typically have weak connections to a high proportion of relative isolates in the user-talk page network |
| Social actions | Particular sets of behaviors which are targeted at other members of a social group – may be measured qualitatively or quantitatively | Vandal fighters exhibit frequent reversions of content, sanctioning of rule violators, efforts to teach and integrate norm violators, and imposition of bans on unrepentant violators |
| Self identification | Individual engages in action which explicitly or implicitly indicates that s/he recognizes that s/he fits this role | Self avowed vandal fighters, membership in vandal fighting-related projects and groups |
| Role Definition | Researcher develops a clear definition of the behaviors and structural signatures that make up a social role – roles may be emergent or recognized within the social group | Vandal fighter, or counter-vandalism editor, is a Wikipedia editor with a particular set of behavioral (reversions, sanctioning, etc) and relationship (weak ties to isolates) attributes |
| Abstract categories, classifications | External researcher classifies social role into theoretically relevant type for research or development purposes | Vandal fighter might be considered an altruist, an enforcer of social norms, or a type of technical editor |

Figure 1. Relationship between roles, actions, and social actors

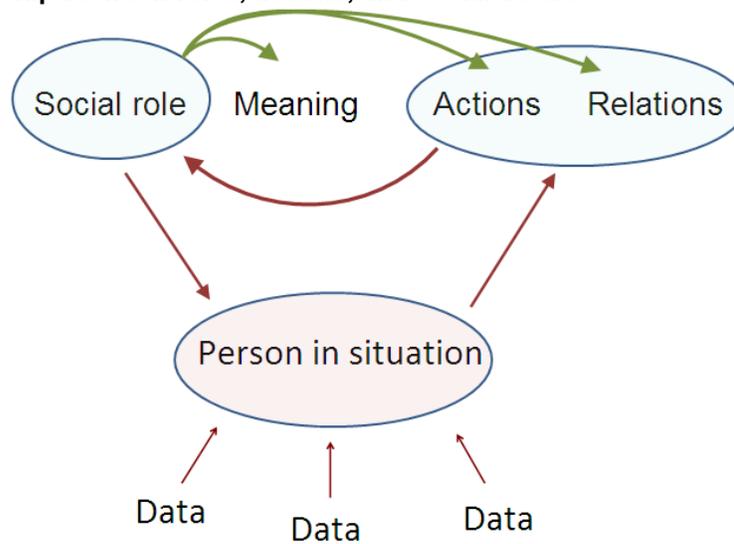


Figure 1 describes the process for identifying roles in online communities and later using structural signatures for predicting the emergence and incidence of a given social role. General qualitative analysis of the content and context of interaction is used to discover behavior that may indicate role use and reveal new role types. Alternatively, quantitative analysis of repeated behaviors or interactions may also uncover indicators for social roles. This initial analytical process entails moving from evidence about persons in situations, to patterns in action and relation in the community, back to situated action. The goal is to iteratively refine the description of a social role. Then, actions relevant to a particular role can be coded via content analysis to establish a ‘ground truth’ for the degree to which a sample of online participants matching the role profile actually engage in that social role during a given time period. This degree of role use can then be compared to patterns of action and social relations from the same time-period. Once a strong correlation between role use and particular structural and behavioral signatures is established, the analysts can use those signatures and other variables to predict variation in role use and to describe a social ecology of roles for given communities. This process allows researchers to predict the existence of roles and whether or not a given set of individuals is engaging in those roles beyond the cases included in the initial content analysis sample. Of course, such extensions are subject to further confirmatory analysis in order to test the generality of the role and the limits of the association between the role and the signatures.

2.3 Utility of definition

Building a catalog of social roles is an important first step towards understanding complex social systems. As a set of social role definitions is developed, these systems can be defined by the interaction of a relatively small set of roles that exist in different proportions, making up a variety of role ecologies. Faced with populations of millions of users, role identification allows these systems to be reduced to the interactions among a relatively small number of roles. With the capability to reduce social media system populations to simplified role ecologies, many practical and theoretical questions can be addressed. Are social media services better served by recruiting more leaders or a relatively larger population of moderate supporters? Are participants playing roles associated with enforcing informal social norms? Do rapidly growing populations overwhelm the ability of existing members to socialize new participants? Which social roles play a disproportionately important function in their environments? Without a model of different roles and their relationships to one another, these questions are difficult to frame, much less empirically evaluate.

In the following, we review social role analysis with a particular focus on methodological approaches and social roles in online communities. The next section presents two examples of the application of structural and content analysis for the identification of roles in two distinct computer-mediated social spaces. The studies presented here rely on content and context to understand social structure. This approach gains significant traction by leveraging computer-mediated

interaction data, providing a low cost, high fidelity source of relationship data and content from naturally occurring social settings. We conclude by advocating for the study of role ecologies to further our understanding of processes in complex social systems, and provide some topics for future research.

3. METHODS FOR SOCIAL ROLE ANALYSIS

Researchers conducting social role analysis have developed two general methodological approaches: interpretive and structural. Interpretive analyses (see e.g. [1] [17]) employ methods such as ethnography, content analysis, and surveys to capture behaviors and relations within groups. This is a prominent method used by anthropologists and sociologists in understanding groups and social systems. Structural analyses (e.g. [7]) employ social network analysis to differentiate individual nodes in a network based on metrics of structure drawn from data on kinship networks, corporate structures, or exchange networks.

Interpretive methods have been used effectively for social role identification [17], understanding individual motivations [31], and observing a population's response to destructive social behavior [18]. While highly useful in identifying and understanding important social roles and the context in which these roles develop, interpretive studies often neglect the macro social structure in which these roles exist. This results in role definitions and findings that are difficult to compare across social settings. Addressing this shortcoming, Nadel [30] proposed explicitly including structural positions, providing the foundations for a 'calculus' of roles. This work was extended into block modeling [52], structural equivalence [24], and role coloration [12]. Others have worked on homomorphism [5] and local roles [25] (for a full review of structural role analysis, see [22]).

These efforts ultimately diverged from the holistic intent of Nadel's proposal and came to equate roles with structural position. Although they contributed a great deal of knowledge about the structure of a network and made it possible to use metrics to point to an individual node or cluster of similar nodes, these studies have moved away from the context and content of the relations. A key example of this is Burt's [7] concept of a broker, a person who links two otherwise disconnected groups of people. Although this concept is theoretically important, it does not effectively distinguish individuals who are engaging in brokering behavior and are socially identified as holding that position from individuals who occupy a structural

position of broker but do not demonstrate brokering behaviors.

A structural conception of social roles provides new insights and is an excellent complement to the interpretive approaches used by ethnographic scholars. Network approaches show that a collection of behaviors can be understood as the performance of a role that leaves an identifiable structural signature [13]. Interpretive approaches provide the contexts and behaviors associated with these structural signatures. Integrating the two approaches by combining the pure structural approach of social network analysis with the behavioral notion of a role [14] makes it possible to identify, based on structural signatures, the social roles that are meaningful at the interaction level. The importance of such roles [16] makes a well-developed hybrid approach to social role analysis highly desirable.

A long tradition in the structural analysis of roles has emphasized that roles entail precise structural positions. While we agree that role use can be revealed through analysis of social structural position, the roles cannot be reduced to a particular position, nor is role use likely to result in exact similarity of positions. For instance, the term 'structural equivalence' refers to situations where two actors have identical ties to and from all other actors in a network ([47]: 356). This concept and its various relaxations are not helpful for identifying roles (as defined here) because they are too precise and too limiting. Instead, social roles are likely to give rise to distinctive tendencies to enter into particular types of relations and tendencies to abstain from other types. For instance, research indicates [50] that experts or 'answer people' will tend to have outward ties to alters who are otherwise isolates in the local community. Furthermore, much discussion of identifying roles in the network analysis tradition (see [47]: 347-502 for a summary), is fixated on the exclusive reliance on structural data to infer roles. In contrast, we contend that efforts to find roles will be most productive when they rely on both structural data and detailed qualitative description of the context and meaning of interaction.

4. ROLES IN SOCIAL MEDIA

We recognize that a holistic approach to role analysis is data intensive. In the past, collecting data on both structure and context was an exceedingly difficult task. Indeed, as a result of data limitations researchers have often been forced to choose a methodological approach that engaged either the structure or the meaning of the behavior that defined each social role. Social media spaces, which include machine-readable traces of social interaction along with highly detailed

behavioral records, represent an unprecedented opportunity to find structurally significant roles and understand their meaning within the social context.

While the costs, quality, and quantity of this class of data are attractive, the greater research significance of this data is that it provides a simplified setting for the study of social roles and the development of social structures. With such data, it is possible to observe the growth and development of social groups and social roles from birth to death, a luxury few other data sets can offer. Furthermore, the abundance of structure and content available through these systems encourages a renewed effort to integrate structural and interpretive approaches to role analysis.

The explosion of online social spaces has produced naturally occurring laboratories for understanding the mechanisms behind the development of informal social organization. In settings such as Usenet and Wikipedia, the actions of any individual within those systems are public. The formation of ties is based on common affiliation with topical areas, not unlike the increased likelihood of friendship forged through common church membership or other shared activities. Some institutions built from internet systems have matured to develop somewhat formal social structures. Wikipedia, for example, has a definable and relatively stable hierarchy of administrators and arbiters. Newcomers who seek to occupy these positions of power and authority must first overcome obstacles that demand significant tenure and history of contribution. These obstacles defend the position of those who have attained them, placing a drag on the rate of social change. Such patterns can be detected in the data contained in the detailed logs of these online systems, providing an opportunity to observe the development of a social system from its earliest formation to its current state. This makes it possible to assess the development of different roles and the relative abundance of a given set of roles over the course of the lifetime of a social system.

The availability of such data, then, represents a golden opportunity to make advances in understanding social roles and role ecologies. Despite the relatively recent availability of such data, numerous studies of social roles and interaction in online communities have appeared (e.g. [32] [42]). However, while some studies have incorporated both interpretive and structural approaches, most studies of social roles have focused strictly on interpretive methods for role discovery (e.g. [17] [18] [31]). Although these studies have identified several important roles in these online groups, there are additional roles that cannot be identified with a strictly interpretive or purely structural approach. Two such roles are described in the following section.

4.1 Example Studies

Combining interpretive and structural methods in social role analysis has great utility. In the examples presented here, investigation into a new social media data set begins with broad qualitative explorations that identify individuals performing interesting roles, after which maps of the structural positions of populations are used to identify network patterns that differentiate users. The next step is to analyze the context of participation and the content of behaviors of the actors whose interactions formed those social network structures. This iterative process moves between content and structure to refine our understanding of social roles and validate the relationship between structural attributes and behaviors. Each of the following cases demonstrates how this combined interpretive and structural method allows for the identification and explication of social roles that either method, in isolation, could miss.

4.1.1. Usenet. Usenet is one of the oldest institutions for social interaction on the Internet. Hundreds of thousands of news servers now connect to one another to globally distribute a collection of a million messages per day organized into more than one hundred thousand newsgroups, which may have a common theme, topic, or simply be the product of a particular collection of people. Collectively, Usenet newsgroups contain messages composed by millions of unique authors.

To discover and describe key social roles within Usenet, Welser *et al.* [50] focused on authors who frequently answer questions in technical support newsgroups. These Answer People provide the bulk of the content in these newsgroups by responding to dozens of questions a day, often more than three hundred days a year. After identifying a few of these Answer People in technical support newsgroups by observing the content they create, others sharing similar structural characteristics were identified. A combination of visualizations of authors' posting behaviors, the posting behaviors of each author's neighbors, and egocentric network graphs allowed newsgroup participants to be coded in terms of their resemblance to an ideal typical Answer Person (for an introduction to structural signatures see [41]). This coding was then correlated with a content analysis of the author's messages to test the structural signature's relationship to answering behavior across several newsgroups and multiple authors.

Just as the Answer Person is a critical component of a technical support newsgroup, the "Discussion Person" and "Discussion Catalyst" play important roles in social support and political discussion newsgroups,

respectively [19]. Discussion People are characterized by frequent reciprocal exchanges with a relatively high number of other participants. Discussion People typically connect to other Discussion People. This social role is the source of most of the discussion content contributed to long threaded conversations. The Discussion Catalyst is responsible for the majority of messages that initiate long threads. The key behavior for these individuals is creating new threads, usually by posting quoted material from external news sources. These root messages then produce long chains of responses as the community picks up the topic and the discussion begins. Interestingly, these individuals often do not participate in the threads they initiate [19].

The egocentric networks presented in Figure 2 show the positional structure for ideal types of the three social roles discussed above. While these differ in important ways, they share similarities that can make them difficult to distinguish, particularly for individuals who have less clearly defined behavioral patterns (for an example, see the Wikipedia roles described below). The context of these relationships, including the newsgroup in which the interaction occurred and the content of the messages exchanged, helps to differentiate one user from another. Answer People and Discussion Catalysts, for example, are distinct social roles that display similar star patterns in their connections to others. Without incorporating the context of technical support newsgroup or political discussion newsgroup, these patterns blur the distinction between the role signatures. It would not have been possible to identify the social roles or understand their significance to the social structures they create without analyzing a combination of structural attributes and contextual clues embedded in the content of the interactions.

4.1.2 Wikipedia. The investigation of social roles has been extended to Wikipedia, the user built, free, online encyclopedia [49]. Wikipedia's goal is to create a set of documents similar to encyclopedia articles. It

differs from a traditional encyclopedia, however, in that any reader may edit or contribute content to any page. The primary tasks of the Wikipedia community focus on constructing and managing these documents. Despite the potential challenges of open access, Wikipedia has cultivated a collective concern for the quality of the resulting articles (see [6]). While Wikipedia has a handful of formally defined roles like administrator and bureaucrat, the largest population of roles are emergent. These informal, collectively defined roles are organized around practices like fighting vandalism, copy editing, enforcing standard formats, welcoming new users, evaluating article quality, and writing tools to help the community [46] [9].

Content analysis of many Wikipedia pages led to the recognition of distinct patterns of behavior. Subsequent structural analysis assessed the similarity between other actors' behaviors and the pattern observed for actors who play particular social roles. This began a process of refining the definitions of the identified roles more cleanly and clearly. The discussion presented here focuses on two critical social roles performed in Wikipedia: Substantive Experts and Technical Editors.

Substantive Experts display extensive knowledge of particular topics, contributing a great deal of content to the subject-area and resolving content-related disputes on the "talk pages" associated with a given Wikipedia content page. Technical Editors enforce the formatting standards of Wikipedia while policing grammar, style, spelling, hyperlink accuracy, and other minor issues. These tasks are related, but many Wikipedia contributors specialize, developing expertise in ways that cause them to distribute contributions differently. These two social roles are of particular interest both because they are crucial to the core task of the encyclopedia and because they possess clearly contrasting structural signatures.

Figure 2. Ego Networks for three roles

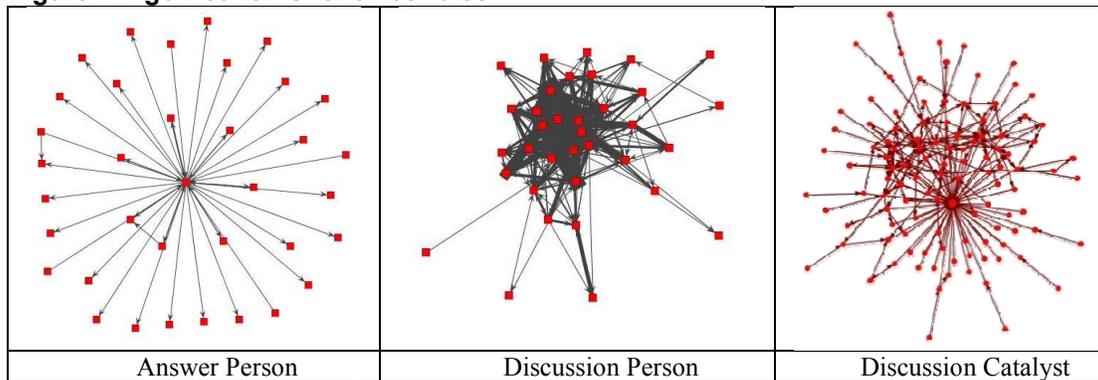
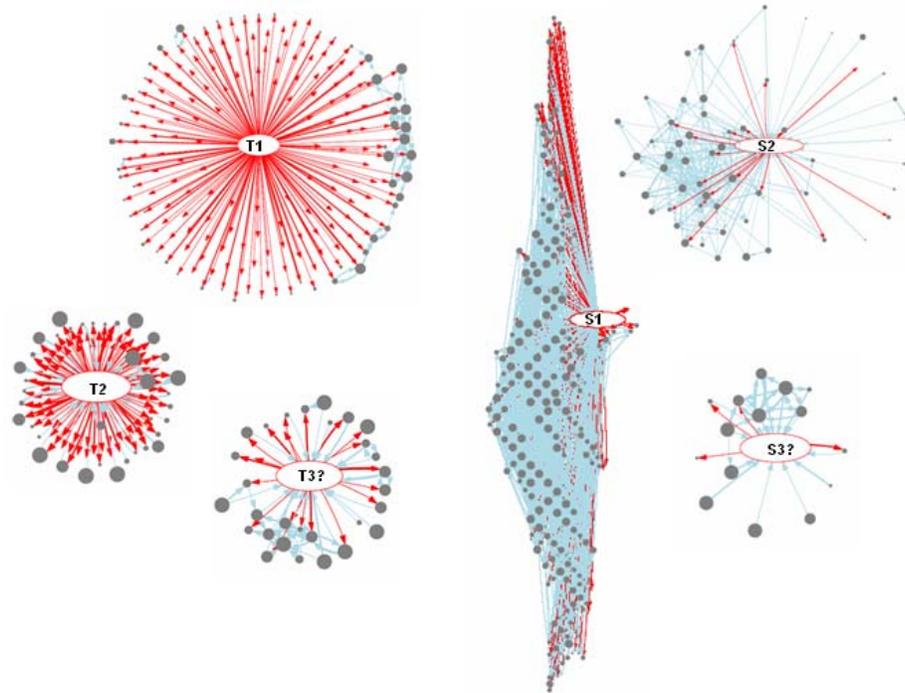


Figure 3. Local Networks for Technical Editors and Substantive Experts



Technical Editors have a repetitive task that is largely independent of the content of any given page. Thus, such Wikipedia contributors should have brief interactions with many other participants, serially moving in and out of relationships with Substantive Experts who have a more focused and durable connection with a smaller set of pages. Substantive Experts, by contrast, are more likely to engage in ongoing conversations with other Substantive Experts as they work to improve pages in their shared areas of expertise. Therefore, the social network structures that surround these social roles may illustrate the distinctions between them. The network structures of Technical Editors should feature outward connections to alters who are otherwise not connected to each other. The performance of the role of Substantive Expert should generate a dense interconnected network.

Figure 3 depicts visualizations for the local social networks surrounding three Technical Editors (T1-T3) and three Substantive Experts (S1-S3). A directed outward tie indicates that an Editor posted a message to another Editor's "user talk page", where conversations associated with each Wikipedia participant are stored. A lighter gray line indicates a reciprocating tie from another Editor. The size of each node indicates the measure of the out-degree from that Editor.

Technical Editors tend to have more outward ties to other Editors who are otherwise unconnected to one another. The illustration of Technical Editor 1 (T1 in

Figure 3) best fits this ideal model, displaying dozens of outward ties and very few connections among those alters. The substantive experts have more variance in the number of alters and the density of their local networks than the technical editors.

The sample networks illustrate the need for contextual analysis in social role discovery. A pure structural approach might confuse some of the technical editors with substantive experts, or vice versa. For instance, T3 and S3 show similar structural features. Analyzing the content of interaction helps us determine whether structural similarity is matched with similarities in behavior and role performance or, alternatively, provide instances where these structural features are not accurate indicators of role performance. This suggests that structural signatures complemented with analysis of content and behavioral contexts can improve our understanding of social roles.

5. ROLE ECOLOGY

Although important, the discovery and cataloging of social roles in order to understand individuals in their social context is not the final goal of this research. Examining social role ecologies, defined by the balance and interaction of roles within a given population, will allow researchers to further our overall understanding of the social world. Understanding these social roles and clearly and consistently defining them

allows us to simplify complex systems and provides a means of comparison between different social settings. With a better understanding of roles, researchers will be able to tackle questions about the relative performance of some populations over others. Furthermore, as researchers develop an understanding of how different role ecologies relate to different social outcomes, these analyses will provide an additional means of studying social dynamics through analysis of the impact of changing role ecologies on social processes.

As [50] demonstrates, social roles are often inherently defined in relational terms; a role only exists in relation to others who are likewise enacting social roles. In a Usenet technical support newsgroup, for example, Question People provide the base material that stimulates Answer People to generate replies. Without a population of questions, an Answer Person cannot exist. In every population in which an Answer Person is present, they remain a small proportion of the population while generating the majority of its content. The bulk of the useful content is thus being contributed by a tiny minority, reflecting the disproportionate importance of Answer People in the newsgroup ecology. Yet each Answer Person requires a large number of Question People in order to support that behavior. Therefore, any single technical support newsgroup can only support a limited population of Answer People proportionate to the active population of Question People. An additional Answer Person in an already saturated newsgroup would only reduce the questions available to other Answer People already present and could lead to the departure of an Answer Person seeking more opportunities to display expertise.

The symbiotic case of Answer and Question People is straightforward; the presence of Answer People is an attraction to Question People while a large population of answerable questions attracts Answer People. Other role ecologies, however, will be less straightforward. A cyclical problem in Usenet, for example, is the invasion of newsgroups by authors seeking to cause social disruption like Flamers, Trolls, and Spammers. When the proportion of messages in a newsgroup from Spammers or Flamers exceeds a threshold, other substantive contributors may leave the group. The withdrawal of key contributors can cause a cascade effect where the signal-to-noise ratio continues to plummet. The result can be the “death” of a newsgroup or a period in which the social space is vacant.

These examples point to the need for a macro orientation to social roles that encompasses both individual behaviors and the whole ecology of social roles within a given social space. As researchers gain a thorough understanding of role ecologies, they can

begin to understand how sudden or gradual shifts in the distributions of roles might impact the overall social system. Understanding the structural positions of these roles and how they are related to one another can further our knowledge of how these social systems function. Such understanding has implications for key theoretical questions from sociology regarding social structure, social roles, and individual agency as well as pragmatic questions regarding community building, the evolution of social groups, and structural qualities of healthy communities.

These research goals are not new and have been thwarted by the challenge of marshalling data of sufficient size and quality. Social scientists have a new opportunity to realize these goals by leveraging recently available data that captures relationships in detail at significant scale. The increased prevalence of computer-mediated social interaction means that social data of this type is increasingly available. As computer-mediated interaction systems evolve and become more widely mobile, they promise to capture a growing swath of social interactions. As more of the social world becomes essentially self-documenting, social roles will increasingly be observable as a function of both social positions, as with early block modeling efforts [52], and content analysis and ethnographic methods.

Computer-mediated interaction data allow researchers to access another layer of analysis not normally available: detailed longitudinal data. It is not just the volume but also the duration of these data sets that will allow us to see more clearly into the social world, grasping processes whose frequencies are too long to be easily recognized by time-limited investigations. These data sets will make it clear that the composition of social roles relative to one another fluctuates over time, revealing either random perturbations and returns to equilibria or the evolution of the social structure. Individual behaviors may evolve from one type of social role to another, or individuals may adopt dramatically different social roles in different social settings. Finally, social data with long time horizons will allow the observation of groups’ responses to environmental changes, such as the appearance of undesirable elements or the departure of core group members. Such analysis could lead to the discovery of strategies that are more successful at maintaining group cohesion and ensuring group survival, information that would be invaluable to anyone trying to understand, develop, or manage a collective process.

6. CONCLUSION

We have argued that research that leverages both structural and semantic dimensions of social roles is consistent with the goals of early network theorists and points to a productive path forward for the structural analysis of social life in computer mediated social spaces. We provide two examples of how researchers can move between interpretive and structural measures in order to refine role definitions. We believe that these strategies will pave the way to a broader study of role ecologies that will help us understand why social outcomes unfold as they do.

Studying how roles are distributed across different social settings will advance research in at least two new directions. The first is community health: when communities rely on key people to enact roles, how does their presence affect the productivity and longevity of the group? The second area concerns the intersection between formal and informal role definitions in organizations. As organizations grow, they increasingly need to formalize roles, yet doing so risks the introduction of roles that are poorly understood and may actually be counterproductive to the goals of the organization. How best can organizations formalize roles in periods of growth and change?

Large-scale structural analyses of roles and role ecologies results in research with implications for other methodological approaches, and even for areas beyond social science. Structural analysis is an excellent complement to ethnographic studies of individual behavior, providing a larger context in which to situate the rich and nuanced analyses these approaches provide. Furthermore, aspects of roles that can only be observed at a macro scale can provide ethnographers with new research questions that can only be answered through qualitative approaches to social science research.

Beyond the academy, understanding social roles can help answer questions with practical applications. Internet search companies seek a means to improve search results and relevance ranking systems. Social search mechanisms are an area of great research interest in the information sciences, particularly with the recent and rapid expansion of online social media. A deeper understanding of social roles is the key to making the transition from the "PageRank" algorithm, famously used by the Google search engine to rank search results, to a "PeopleRank" model that considers the social role the author of a piece of content plays [54].

Finally, social role analysis provides a framework for the integration of disparate areas of study. By extending social network definitions of roles and seeking to improve our understanding of how to measure social interaction, social roles research helps

develop baseline measurements that can be used to compare and contrast entities across a variety of fields. As more of social life is conducted in or through computation, we foresee a future in which social sciences gain new traction on long standing questions and an unprecedented depth of insight into large-scale patterns of human association.

7. REFERENCES

- [1] Anderson, Elijah *Code of the Street: Decency, Violence, and the Moral Life of the Inner City*. New York: W.W. Norton. (1999).
- [2] Baker, W.E. and Faulkner, R. R. "Role as Resource in the Hollywood Film Industry." *American Journal of Sociology* 97:279-309. (1991).
- [3] Biddle, B. J. *Role theory: Expectations, identities, and behaviors*. New York: Academic Press. (1979).
- [4] Blau, J.R. and Goodman, N. *Social Roles and Social Institutions: Essays in Honor of Rose Laub Coser*. Boulder: Westview Press. (1991).
- [5] Bonacich, Phil. "What is a Homomorphism?" In *Research Methods in Social Network Analysis*, Freeman, Linton C., White, Douglas R., and Romney, A. K., Eds. George Mason University Press. Fairfax, Virginia. (1989).
- [6] Bryant, Susan, Andrea Forte and Amy Bruckman "Becoming Wikipedian: Transformation of Participation in a Collaborative Online Encyclopedia." *Proceedings of GROUP: International Conference on Supporting Group Work*, Sanibel Island, FL. pp 1-10. (2005).
- [7] Burt, R. *Brokerage and Closure: An Introduction to Social Capital*. Oxford: Oxford University Press. (2005).
- [8] Callero, P. L. "From Role-Playing to Role-Using: Understanding Role as Resource." *Social Psychology Quarterly* 57 (1994) 228-243.
- [9] Cosley, D., Frankowski, D., Terveen, L., & Riedl, J. (2007). SuggestBot: using intelligent task routing to help people find work in wikipedia." *Proceedings of the 12th international conference on Intelligent user interfaces*. ACM Honolulu, Hawaii, USA
- [10] Donahue, E. M., Robins, R. W., Roberts, B. W., & John, O. P. "The divided self: Concurrent and longitudinal effects of psychological adjustment and social roles on self-concept differentiation". *Journal of Personality and Social Psychology*. 64(1993) 834-846.
- [11] Donath, J. "Identity and Deception in the Virtual Community." *Communities in Cyberspace*. Ed. M. Smith and P. Kollock. London:Routledge. (1998).
- [12] Everett, Martin G., and Borgatti, Steve. "Role Colouring a Graph". *Mathematical Social Sciences* 21(1991)183-8.
- [13] Faust, K. and Skvoretz, J. "Comparing Networks Across Space and Time, Size and Species". *Sociological Methodology* 32 (2002) 267-299.
- [14] Fiske, S.T., and Taylor, S.E. *Social Cognition*. New York: McGraw-Hill (1991)
- [15] Gladwell, M. *The Tipping Point: How Little Things Can Make a Big Difference*. Little, Brown and Company. (2000).
- [16] Goffman, E. *The presentation of self in everyday life*. New York: Doubleday. (1959).

- [17] Golder, S. and Donath, J. "Social Roles in Electronic Communities". Presented at *Association of Internet Researchers IR 5.0*. September 19-22, 2004, Brighton, England. (2004).
- [18] Herring, S., Job-Sluder, K., Scheckler, R., and Barab, S. "Searching for Safety Online: Managing Trolling in a Feminist Forum." *The Information Society*. 18(5). (2002).
- [19] Himelboim, I., Fisher, D., Gleave, E., and Smith, M. "Reply Magnets in Online Political Discussions: Analysis of Six Months of Discussion in 20 Usenet Political Newsgroups". Paper presented at the Association for Education in Journalism and Mass Communication 58th Annual Convention. Washington D.C., August 9-12. (2007).
- [20] Jackson, L.A, and Sullivan, L.A, Perceptions of multiple role participants, *Social Psychology Quarterly*, 53:3 (1990) 274-282.
- [21] Lento, T., Welser, H. T., Gu, L. and Smith, M. A. "The Ties that Blog: Examining the Relationship Between Social Ties and Continued Participation in the Wallop Weblogging System". Paper presented at Workshop on the Weblogging Ecosystem, 15th International World Wide Web Conference, Edinburgh May 23-26. (2006).
- [22] Lerner, J. "Role Assignments" pp. 216-252. in *Network Analysis: Methodological Foundations*. Ed. Brandes, U., and Erlebach, T. Springer-Verlag Berlin. (2005).
- [23] Linton, R. *The Study of Man*. New York: D. Appleton-Century. (1936).
- [24] Lorrain, F., and White, H.C. "Structural equivalence of individuals in social networks". *Journal of Mathematical Sociology*. 1, (1971). 49-80.
- [25] Mandel, Michael J. "Local Roles and Social Networks". *American Sociological Review*. 48:3 (1983) 376-386.
- [26] McCall, G. J., & Simmons, J. L. *Identities and interactions*. New York: Free Press. (1978).
- [27] McPherson, Miller, Smith-Lovin, Lynn, and Cook, James M. "Birds of a Feather: Homophily in Social Networks". *American Review of Sociology*. 27 (2001). 415-44.
- [28] Mead, G. H. *Mind, Self & Society: From the Standpoint of a Social Behaviorist*. Ed. C. W. Morris. Chicago: University of Chicago Press. (1934).
- [29] Merton, R.K. "Continuities in the Theory of Reference Groups and Social Structure." Pp. 422-440 in *Social Theory and Social Structure*, ed. R.K. Merton. New York: Free Press. (1968).
- [30] Nadel, S. F., *The Theory of Social Structure*. Cohen & West. London. (1957)
- [31] Nonnecke, B. and Preece, J. "Lurker Demographics: Counting the Silent" *Proceedings of CHI 2000*.
- [32] Nonnecke, B. and Preece, J. "Why Lurkers Lurk". *Americas Conference on Information Systems*. (2001).
- [33] Park, R. E. "Behind Our Masks." *Survey Graphic* 56 (May) (1926) 135.
- [34] Parsons, T. *Essays in Sociological Theory, Pure and Applied*. Glencoe, Illinois: The Free Press. (1949).
- [35] Parsons, T. *The Social System*. Glencoe, Illinois: Free Press. (1951).
- [36] Parsons, T. *Action Theory and the Human Condition*. New York: Free Press. (1978).
- [37] Reitzes, D.C. and Mutran, E. J. "Self-Concept as the Organization of Roles: Importance, Centrality, and Balance." *Sociological Quarterly*. Vol. 43. Issue 4. (2002). pp647-667.
- [38] Sarbin, T. R., & Allen, V. L. "Role Theory". In *Handbook of social psychology*, Volume I (pp. 488-567). Reading, MA: Addison-Wesley. (1969).
- [39] Shneiderman, Ben, "Computer Science: Science 2.0", *Science* 319(5868): 1349-50. (2008)
- [40] Simmel, G. *Soziologie: Untersuchungen uber die Formen der Vergesellschaftung*. Leipzig: Duncker and Humblot. (1908).
- [41] Skvoretz, J., and Faust, K. "Relations, Species, and Network Structure". *Journal of Social Structure*. Vol. 3 No. 3. (2002).
- [42] Smith, M.A. and Kollock, P. (Eds.) *Communities in Cyberspace*, London: Routledge. (1999)
- [43] Stryker, S. *Symbolic Interactionism*. Menlo Park, CA: Benjamin Cummings. (1987).
- [44] Turner, T. C., Smith M.A., Fisher, D., Welser H. T. "Picturing Usenet: Mapping Computer-Mediated Collective Action." *Journal of Computer Mediated Communication* 10(4). (2005).
- [45] Viegas, F.B. and Smith, M.A. "Newsgroup Crowds and Authorlines: Visualizing the activity of individuals in conversational cyberspaces". *Proceedings of the 37th Hawaii International Conference on System Sciences*. Los Alamitos: IEEE Press. (2004).
- [46] Viegas, F.B., Wattenberg, M., Kriss, J, and van Ham, F. "Talk Before You Type: Coordination in Wikipedia". *Proceedings of the 40th Annual Hawaii International Conference on System Sciences*, p78 (2007).
- [47] Wasserman, S., Faust, K. *Social Network Analysis*. Cambridge: Cambridge University Press. (1994).
- [48] Weber, S. *The Success of Open Source*. Cambridge MA: Harvard University Press. (2004).
- [49] Welser, H. T., Cosley, D., Kossinets, G., Lin, A., Dokshin, F., Gay, G., and Smith, M.A. "Finding Social Roles in Wikipedia". ASA, Boston. (2008).
- [50] Welser, H. T., Gleave, E., Fisher, D., and Smith, M.A. "Visualizing the Structural Signature of Social Roles in Online Discussion Groups". *Journal of Social Structure* 8:2. (2007).
- [51] Welser, H. T., Smith, M.A., Gleave, E., and Fisher, D. "Distilling Digital Traces: Computational Social Science approaches to studying the internet" *Handbook of Online Research Methods*. Eds Fielding, N., Lee, R.M., and Blank, G. (2008)
- [52] White, H. C., Boorman, S. A., and Breiger, R. L. "Social Structure from Multiple Networks. I. Blockmodels of Roles and Positions", *American Journal of Sociology*. 81:4 (1976).
- [53] Winship, C.. "Thoughts about Roles and Relations: An Old Document Revisited." *Social Networks* 10 (1988) 209-231.
- [54] Xi, W., Lind, J., and Brill, E. "Learning Effective Ranking Functions for Newsgroup Search", SIGIR '04, July 25-29, 2004, Sheffield, UK.