Exploring Equivalence Metrics to Analyze Behavior in a Naked Mole Rat Colony

Susan P. Imberman, Ph.D.[4,5], Michael E. Kress, Ph.D.[3,4,5]. Dan P. McCloskey Ph.D.[1,2]

1.Dept. Psychology 2. CSI/IBR Center for Developmental Neuroscience, 3. Vice President of Technology Systems, 4. Dept. Computer Science College of Staten Island, 2800 Victory Blvd, Staten Island, NY 10314, 5. Doctoral Program Computer Science the Graduate Center at CUNY,

Correspondence to: susan.imberman@csi.cuny.edu

Abstract

Many statistical and data mining techniques have been used to analyze the deluge of data generated by computerized, sensing devices. Behavioral psychologists traditionally have relied on "low-tech" methodologies for observing animal behavior in the wild and the laboratory. These methods are time intensive and laborious. When the observed animal is a colony animal, with many individuals to observe, traditional methods fail We inject RFID passive transponders under the skin of our study animal, the Naked Mole-rat (NMR). RFID readers are placed throughout the housing environment, allowing us to track animal movements as they move through these areas, with sub-second resolution for long periods of time. This methodology generates huge amounts of data requiring Big Data analytical techniques. In this paper, we investigate equivalence metrics, specifically the Pearson Correlation Coefficient and Hamming Distance, to analyze behavior changes in the social network structure of a naked mole rat colony. Our results showed that a Pearson Correlation was sufficient to detect equipment error and Hamming Distance could detect changes in colony behavior.

Background and Motivation

Traditional tools used to measure behavior of laboratory animals become problematic when large groups of animals need to be measured simultaneously. Our recent work has examined social behavior using network tools in a highly specialized laboratory-housed rodent, the African naked mole-rat (Heterocephalus Glaber). African mole-rats are unique among rodents, indeed among all mammals, because they participate in a cooperative breeding eusocial lifestyle. Naked mole-rats rely on a queen to be solely responsible for breeding the entire colony. The lack of reproductive ability, and close kin relationships among colony members, allows the naked mole-rat to maintain the largest colony size of any mammal with up to 300 members in wild colonies, and up to 100 animals in captive colonies.

Naked mole-rats are quickly becoming an important model in biomedical research. The naked mole-rat genome has more genes in common to humans than traditional laboratory rodents (rats and mice; Kim, et al. 2011), yet they demonstrate cancer and stroke resistance (Larson & Park, 2009) and extreme longevity, living over 30 years (Buffenstein, et al. 2008). These animals thrive in captivity and allow for the study of their cooperative and social behavior. Understanding the brain mechanisms driving the social behavior in the naked mole-rat may help us to further understand human social behavior and social behavior disorders such as autism and schizophrenia.

We have presented the first colony-wide analysis of activity patterns and social behaviors of naked mole-rats using customized radio frequency identification (RFID) based tracking system (McCloskey et al. 2011). Using this system, we can track all movements of all animals as they move through all areas of their laboratory habitat with sub-second resolution for long periods of time. The event-based dataset generated by this type of system allows for the simultaneous study of individual, sub group, and whole group characteristics. Our initial studies in this area were focused on identifying social behaviors of animals in a colony based on the premise that animal interaction can be determined by measuring how much time is spent in the same area of the housing environment. We found that using a combination of data mining approaches, including adjacency matrix sampling, principal component analysis, and frequent pattern mining, we can successfully identify characteristics of the social network (McCloskey et al. 2011).

We have also used the RFID data to demonstrate that naked mole-rats, like all animals, have a periodicity to their behavior. The circadian cycle of activity is roughly 24 hours in the presence of external cues such as light and interaction with experimenters and animal care staff.

When adjacency matrix sampling and frequent pattern mining were analyzed, in relation to time of day, we found that the density of social networks and the complexity of behavior patterns varied across the circadian rhythm (Imberman et al. 2012).

Our current goal is to establish a long-term recording system so that the state of social behavior can be measured continually, and any deviation form the baseline level of behavior can be readily detected. In order for long-term recording to be successful, it would be helpful to have an "on-the-fly" monitoring system to detect deviations in the behavior pattern of one animal, a small group of animals, or the entire colony, while monitoring the entire network in real time. The present study simulates this deviation by analyzing data generated by our RFID sensors. Our first study looks at a synthetic perturbation of a 1 hour time window by keeping some subset of RFID transponders in a single location for one hourly time window. This was analyzed in multiple simulated datasets ranging from one animal to the entire colony, and measured in three time windows of different baseline activity. Two approaches were compared for their ability to detect changes in the resulting adjacency matrices of network activity. The first, approach was to perform Monte Carlo simulations of the perturbed dataset and conduct a Pearson Correlation between the original and the perturbed network data. The second, approach was to perform a Quadratic Assignment Procedure (QAP) and measure changes in the Pearson Correlation Coefficient when comparing the original and perturbed datasets. In our second study we looked at multiple windows of activity and compared an active window with all other windows over a 27 hour period. The thesis here was to try to detect different behavior, such as sleep, from other activity. These approaches are promising for incorporation of an "on the fly" monitoring system into the laboratory setup.

Methods and Technical Solutions

For our initial study, 33 NMRs were housed in standard mouse tub cages connected by over 7 meters of clear polycarbonate tubing (50.8 mm inner diameter). Each NMR was implanted subcutaneously with a Trovan Unique radio frequency identification transponder, referred to henceforth as a "chip" (transponder size 11.5 x 2.2mm; MicrochipID Lake Zurich, IL). Stationary circular RFID reader antennae, (Trovan LID 650 readers (MicrochipID, Lake Zurich, 100 mm inner diameter) were placed around the polycarbonate tubing at multiple locations and connected to a computer so as to facilitate data collection. Each time a tagged NMR passed through a reader, a text file was updated with the animal ID (unique 10 digit alphanumeric code), time of entry, and reader number (1-14). Data preprocessing of 24 hours of data entries was organized into a state matrix identifying the last known location for each animal, for each event. This methodology produces a large amount of data, with approximately 4 events per second or 15,000 events recorded per hour.

From the state matrix we calculate a Total Adjacency Matrix, which shows the total number of times each dyad (pair) of animals collocated in that time window (Table 1). Hence in Table 1, NMR 2 collocated with NMR4 five times in this time window.

Table 1 Sample Total Adjacency Matrix

	NMR1	NMR2	NMR3	NMR4	NMR5	NMR6
NMR1	0	0	0	0	0	0
NMR2	0	0	0	5	2	0
NMR3	0	0	0	0	0	2
NMR4	0	5	0	0	0	0
NMR5	0	2	0	0	0	0
NMR6	0	0	2	0	0	0

Previously, we were able to identify, using cluster analysis, graphical models, and frequent pattern mining, colony wide behavior patterns. Using traditional social network central graph analysis and frequent pattern mining with Apriori at 50% support, (Agrawal et. al. 1994), we were able to find equivalent sociograms. In (Imberman et. al.2012) we showed that by using these methods, in addition to principal components analysis, activity level over a 24 hour period can be quantified and visualized. High levels of activity showed a different picture than low levels of activity. Using these methods, we can identify time windows of one hour long duration showing different levels of activity. For example, in the initial data set, Window 1 exhibited low activity, window 7 moderate activity, and window 14 high activity. Figure 1 shows histograms of these windows, visualizing the various activity levels as a function of animal frequency at a reader.

Simulated Chip Failure

Given the methods described, the question becomes, "How large are the changes in network behavior, and thus changes in social behavior, before we can reliably detect them?" Changes in network behavior can be attributed to several factors. One is that an event occurs within the colony to warrant a change in behavior. An example of this can be the birth of pups or the death of a queen. Changes in network behavior can also be attributed to equipment errors, which in our case would be either a reader failure or a chip failure causing one, the other, or both to go offline.

For the initial study, we focused on equipment errors. In our laboratory setup, readers were arranged linearly. To create the state matrix, and from that the total adjacency matrix, we assume an animal has not moved from its last

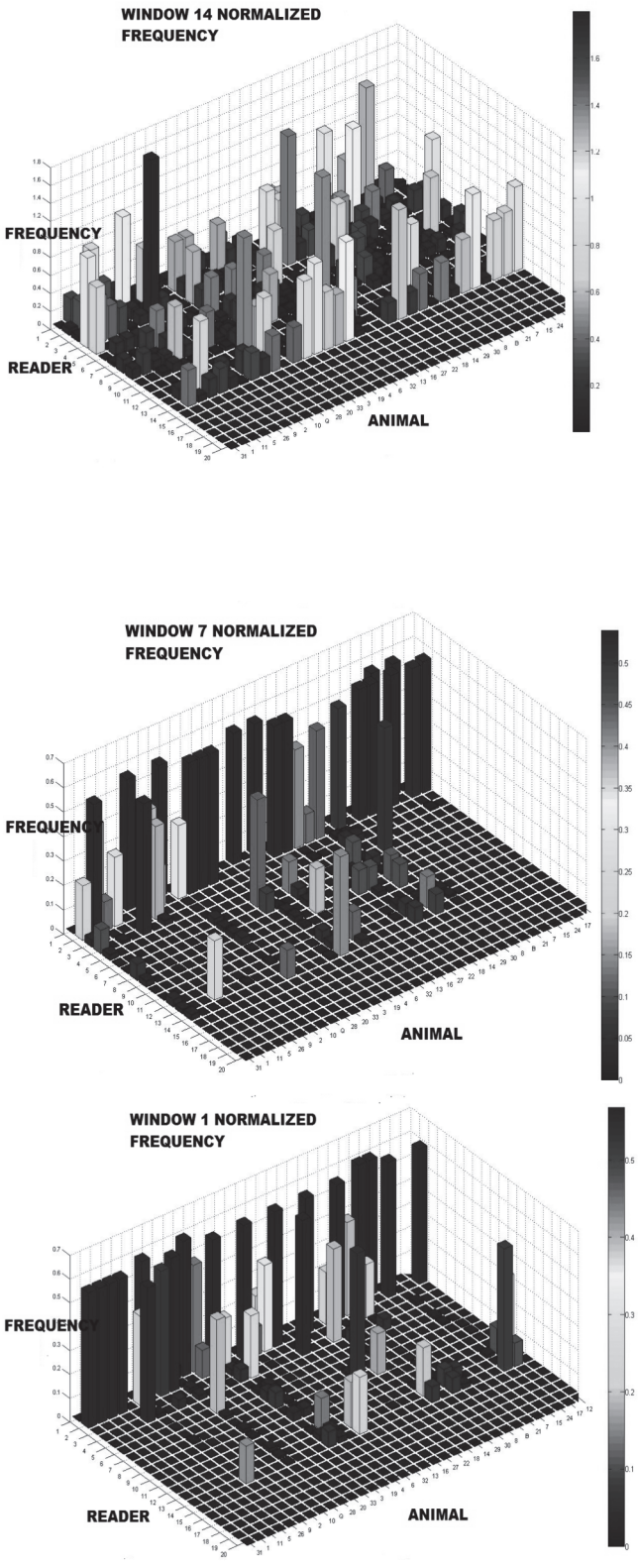


Figure 1 Histograms showing frequency of localization for a given reader in a given location. for each NMR

position until it triggers a new reader. A reader failure results in an NMR being counted as not having changed position until another reader, further along in the sequence, is triggered. The behavior of interest is animal interaction, and whether groups of animals remain together. As long as the readers are spaced close enough together, reader failure would have a minimal impact on these results.

On the other hand, should an RFID chip fail, the animal interaction data can be affected. The ability to detect such an event is the basis for our first empirical study. This type of change in behavioral data would occur in the case where the implanted chip becomes nonfunctional or falls out. In addition, a chip can appear to be nonfunctional if the animal itself suddenly becomes stationary. This could occur, for example, if the animal becomes either sick or deceased, or if a queen is giving birth. A nonfunctional picture in the data means that the animal appears to stay at a particular location (reader) for an extended period of time, or for this study, a full hour. The question remains, how sensitive is our network, and hence our data analysis, to failure in RFID chips, which we will refer to as chip error. By partitioning our data into windows of hour long time periods, we can look at colony activity at different time periods during the day. We chose one day's worth of data, gathered early in our collection process for this initial study. As previously mentioned, windows 1, 7, and 14 respectively represent low, medium and high amounts of colony activity (Figure 1). Hence, another question arises: "Is there a delta in which chip failure has a minimal effect on how we categorize colony behavior?"

Our initial analysis involved a Monte Carlo simulation of chip error. NMR were randomly selected for chip failure, with the assumption that this failure happened at the start of a time window. A base case was created (no error) for each of the time windows 1, 7 and 14. Error cases were created for failure of one chip/NMR up to all 33 chips. Each error case was done for 600 replications and the Pearson Correlation Coefficient was calculated for the comparison between the no error case to the case for each replication. Figure 2 is a graph of the results over the 3 windows, W1, W7 and W14, with the results of the Pearson Correlation averaged over the 600 trials.

As expected, the base run case has a Pearson Correlation equal to 1 and the case of 33 (max number of animals) is a number between 0 and 1 for each of the windows. From the graph one can see that the curve of Pearson Correlation Coefficient versus the number of chip errors is monotonic. The rate at which the curves degrade differs with colony activity. Chip failure seems to show a more marked difference as the activity level of the colony increases. This is not surprising since a chip failure is tantamount to an NMR remaining stationary. When comparing to the base case, which is an active state, one would expect to see a larger

difference between a "totally stationary" set of animals versus an active set.

Looking at the results another way, we want to make sure that the difference in correlation due to chip error is not the same difference we would see in a randomly generated sample. The relationships between mole-rats seen in the total adjacency matrix are not independent. Therefore there is a dependency between values within a row, and values within a column. When this is the case, a Quadratic Assignment Procedure (QAP) is run to obtain a random matrix where the dependencies

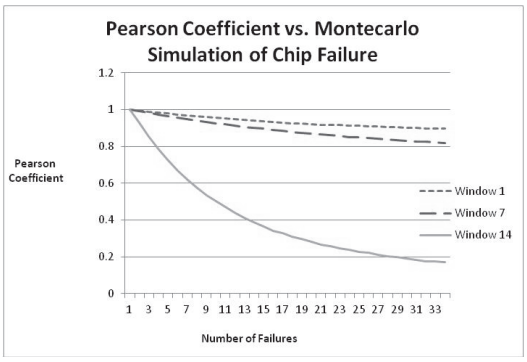


Figure 2 Results of Monte Carlo Simulation

within a row and column are preserved, but the dependencies between compared networks are broken. (Krackhardt, D. 1987; Krackhardt, D. 1988, Borgatti et. al.). If we run Pearson's Correlation on each of these permuted matrices, we can determine the significance of the correlation between the model (chip error) and the base run due to random effects. Figure 3 shows a graph of the results of the QAP Correlation Procedure in UCINET1 for Chip Error of 1 to 10 chips.

Detecting Behavior Changes

Histogram analysis has shown that the colony has periods of varying levels of activity. As mentioned, we can categorize these as high levels of activity, medium levels of activity and low levels of activity. Low activity levels are usually when the NMR sleep. The purpose of our second related study was to determine if we can use an equivalence metric to quantify these behavior periods, or, in essence, when do animals sleep? The data used for this study was a 28 hour period broken into 1 hour windows. The first hour was an active period. Figure 4 shows histograms of three time windows in this 28 hour period. As you can see the activity changes from high to low (most animals stay at readers 19 and 20) and back to high. Window 17 is an example of a sleep period.

[1] https://sites.google.com/site/ucinetsoftware/home

Total adjacency matrices were calculated for each window. Since each window can have a different number of readings, the matrices were normalized using a threshold. The threshold matrix was created by entering a 1 if the dyad existed at least the calculated percentage of total readings in that time window. Thresholds were set at 50 percent and 75 percent.

There are several equivalence measures used in social network analysis, the Pearson Correlation Coefficient being one of them. Buoyed by our success with the synthetic dad, we applied the Pearson Correlation coefficient to the threshold matrices. The matrix for Window 1 was compared to each successive matrix for Windows 2 – 28. One would expect some difference between an active matrix and a "sleepy" one. Figure 5 is a graph of the results of the Pearson Correlation.

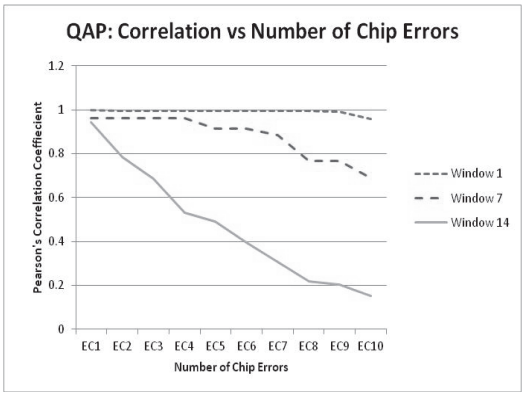


Figure 3 Results of QAP Simulation

As we can see, the sleep behavior is not picked up by the Pearson correlation metric. According to Hanneman, "The correlation measure of similarity is particularly useful when the data on ties are "valued," that is, tell us about the strength and direction of association, rather than simple presence or absence." Since our threshold matrix indicates presence or absence of a significant dyadic collocation, the Pearson is not a good metric to indicate similarity/dissimilarity of the underlying social network for these matrices.

Hamming distance is another equivalence metric used to show dissimilarity between networks. The Hamming distance calculates the number of dyads that differ. Hence the larger the number, the more the matrices differ. Figure 6 shows the results of Hamming Distance applied to both the 50 percent thresholded matrix and the 75 percent thresholded matrix. The Hamming distance metric is able to distinguish between active behavior and sleep behavior. The values peak around window 17 which corresponds to the picture painted by the histograms. In comparing the 50 percent threshold matrix's results with the 75 percent matrix, one sees less variation. This can be due to the fact

that animals that collocate frequently, do this consistently. Interestingly, the results also show that the some animals tend to sleep less than others. This behavior was not previously known.

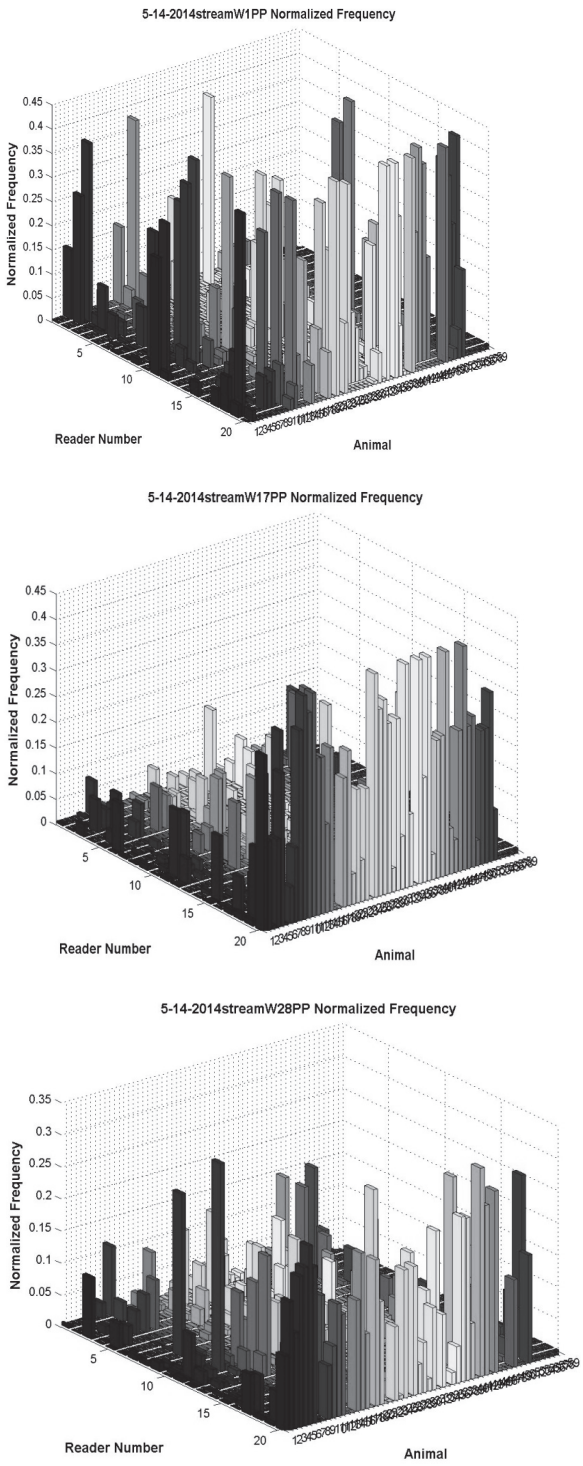


Figure 4 Histograms of Windows 1, 17 and 28

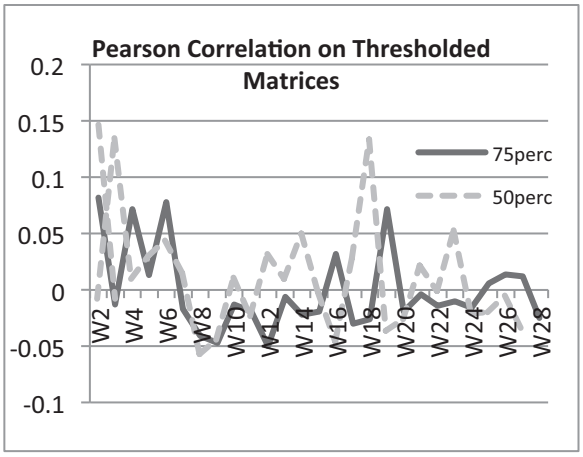
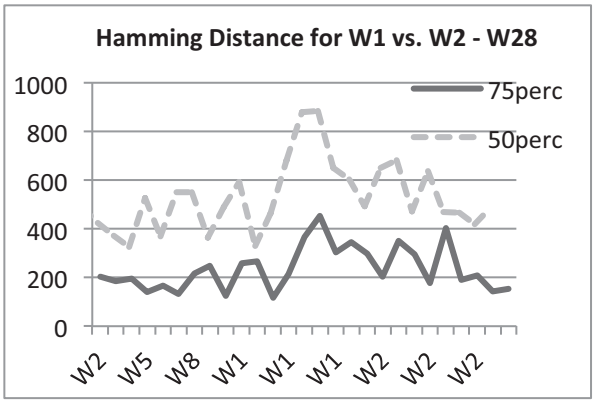


Figure 5 Results of Pearson Correlation



Figurē 6 Results of Hamming Distance

Significance and Impact

In this study we looked at two equivalence measures, the Pearson Correlation Coefficient and Hamming Distance. The Pearson measured the ability to detect a change in the network by analyzing the correlation between simulated perturbed datasets and the original dataset using two approaches. Both the Monte Carlo simulation and the Quadratic Assignment Procedure approach were capable of showing changes in the network pattern, and in both cases the sensitivity was directly related to the time of day and the number of animals affected. Hamming Distance was better than the Pearson Correlation Coefficient at differentiating between behavior patterns from different time periods.

Our previous analysis (Imberman, et al. 2012) demonstrated that time of day is a key factor in determining network behavior in the naked mole-rat. During the night period animals tend to sleep together in the nest area and stay huddled for warmth, which results in a high density social network. In the morning and evening hours networks remain relatively dense although sleeping has sub-

sided in most animals and movement around the colony has increased.

In the current study, time of day was an important factor in the ability to detect network change. In the chip error study, it is clear from both the Monte Carlo and QAP analyses that animals remaining still during window 14 (when most animals are moving around) are more easily detected than during window 1 (when most animals are sleeping). Of course, we would predict the opposite result if the perturbation of the dataset was the simulated movement of an animal through the colony. Therefore, a useful monitoring system will need to detect erratic behavior patterns relative to the baseline pattern for the colony at that time.

As predicted, the analyses were also sensitive to the number of animals displaying an altered behavior pattern. When the simulated chip loss was limited to one animal, the correlation to the original adjacency matrix remained near 1, independent of the time of day or method used. However, as the number of animals simulated as remaining still (or losing chips) increased, the correlation coefficient dropped off quickly. These data suggest that when a monitoring system is used on an hourly basis, there is a critical mass of animals required with changed behavior patterns to reliably detect a change in the overall network.

Pearson Correlation Coefficient did not work well to capture the differences in time of day behavior. Hamming distance was a much better measure of this. An important question was what equivalence measure can be used as an "on-the-fly" method for detecting colony behavior changes. Our findings show that in detecting equipment errors the Monte Carlo simulation approach provides an accurate reflection of the effects of network change, by virtue of the smoothness of the curve in all scenarios. One would expect nearly linear reductions in correlation coefficients with incremental increases in network perturbations, such as the number of animals simulated to remain still. However, the QAP analysis, which requires significantly less computation, seems adequate in approximating the network changes shown with the Monte Carlo approach. Hamming Distance may prove more useful in detecting global behavior changes. In our goal to monitor whole colonies of naked mole-rats over long periods of time, network comparisons using all three approaches may provide different kinds of feedback on colony behavior. This can mean that researchers can be notified in a timely manner as to possible significant events in colony behavior.

In summary, this study of a network equivalence metrics shows several viable options for flagging behavioral events. While it appears unlikely that we would be able to detect subtle changes in the behavior pattern of one animal during one hour of analysis, it is possible to identify changes in the behavior of a small group of animals, and even the colony as a whole. Our sensitivity is directly influenced by the baseline behavior pattern of the entire colony. Therefore, an accurate monitoring system should be able to adjust the sensitivity based on the variability in the baseline behavior. We find the Quadratic Assignment Procedure with Pearson Correlation to be nearly as sensitive as Monte Carlo simulation with Pearson Correlation. Also, Hamming Distance is a better metric than the Pearson Correlation Coefficient for identifying changes in colony behavior.

Acknowledgments

Our thanks to Scott Imberman for suggesting the use of Monte Carlo methods. This research was supported, in part, under National Science Foundation Grants CNS-0958379 and CNS-0855217, PSC-CUNY grant 65289-00-43, the Dormitory Authority of the State of New York, and the City University of New York High Performance Computing Center.

References

Agrawal, R., and Srikant, R. 1994. Fast Algorithms for Mining Association Rules. IBM Res. Rep. RJ9839. IBM Almaden.

Borgatti, S., Everett, M. Jonson, J., Analyzing Social Networks, Sage Publishing 2013.

Buffenstein, R., 2008. Negligible senescence in the longest living rodent, the naked mole-rat: Insights from a successfully aging species. Journal of Comparative Physiology and Behavior, Biochemical, Systemic, and Environmental Physiology, 178(4), 439-445.

Hanneman, Robert A., Introduction to Social Network Methods, online text, http://faculty.ucr.edu/~hanneman/nettext/C13_%20Structural_Equivalence.html#toc

Hubert, L. J. & Schultz, J. (1976). Quadratic assignment as a general data analysis strategy. British Journal of Mathematical and Statistical Psychology, 29, 190-241.

Imberman, S.; Kress, M.; McCloskey, D.. Using Frequent Pattern Mining To Identify Behaviors In A Naked Mole-rat Colony. in Florida Artificial Intelligence Research Society Conference, North America, May. 2012.

Kim, EB, et al. (2011) Genome sequencing reveals insights into physiology and longevity of the naked mole rat. Nature 479:223-227.

Krackhard, D. 1987. QAP Partialling as a Test of Spuriousness. Social Networks, 9: 171 - 186.

Krackhardt, D. 1988 Predicting with Networks: A Multiple Regression approach to Analyzing Dyadic Data. Social Networks, December 10(4) 359 – 381

Larson J, Park TJ (2009) Extreme hypoxia tolerance of naked mole-rat brain. NeuroReport 20:1634-1637.

McCloskey, D. P., Kress , M. E., Imberman, S. P., Kushnir, I, and Briffa-Mirabella, S.. 2011. From market baskets to mole rats: using data mining techniques to analyze RFID data describing laboratory animal behavior. in Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '11). ACM, New York, NY, USA, 301-306