Effect of Power on Trust and Autonomy

Prapulla Poli
Computer Science & Computer Engineering Department
Engineering Hall, Room 313
Fayetteville, AR 72701
+01-479-575-2420
ppoli@uark.edu

Henry Hexmoor
Computer Science & Computer Engineering Department
Engineering Hall, Room 313
Fayetteville, AR 72701
+01-479-575-2420
hexmoor@uark.edu

ABSTRACT
This paper presents a simple model of trust and autonomy and explores the role of power as social influence on trust and autonomy. We present a quantitative measure we term power homogeneity that quantifies profiles of power assignment in a group. Experimental results illustrate changes in trust and autonomy with groups at different power homogeneity. We also show that our model of autonomy predicts commitment levels.

1. INTRODUCTION
Multi-agent systems are useful in modeling organizations where there are varying group sizes and relative power among individuals. Studies of sociality in multiagency strive to provide individual agent decision making with abilities to account for interaction and group awareness. Autonomy, trust, and power are among the social notions that are used to qualify interaction and awareness. Differing power levels and group sizes change trust and autonomy among agents. The main position of this paper is that different levels of power contribute to different trust and autonomy values, which are used as part of the delegation decision. We compare average autonomy and average trust in groups with different power levels. Subgroups of agents with low and high power experience high autonomy and trust than those agents with medium power in the group. We extend the principle of social entropy given in [2] to power homogeneity and use it in our comparisons. The power homogeneity of the agents were used to experiment with two conditions on commitment among agents to delegate a task vary with varying power homogeneity levels and whether autonomy can be used to predict the commitment trends.

Social power in general can take on different forms such as economic power, knowledge as power, personal power, physical power, situational power, etc. Our notion of power focuses on relative power. In this paper we consider power as relative influence among agents when actions are delegated. A form of power as influence is given in [3]. Also models of power and trust in organizations are found in [5]. We treat an agent's influence value to be uniform with respect to all agents. This type of power is akin to ranks in an organization that creates hierarchy among agents. Power levels distinguish agents from one another.

Trust between agents depends on many parameters including competency, histories, reciprocity, benevolence, culture, and reputation [4]. Interpersonal trust is also a function of familiarity and social ties. Interpersonal trust, which is the focus of this paper, is similar to institutional trust. Institutional trust is the trust that exists among individuals due to their participation in social norms and values of various institutions they are members of. Usually, trust levels accumulate and diminish gradually unless there are radical changes in agent attitude toward one another, such as major changes in benevolence [1]. Another conceptualization is that, trust is not a precursor to delegation but one between collaborating individuals who communicate. Trust is in the degree of belief in validity of messages. In this notion of trust, capability, benevolence and exchanges of trustee is not in question but the agents' interaction with different power levels is considered. In summary, we suggest that agent X's trust in agent Y about task T (we will denote that by Trust (X, Y, T)), is partly a function of agent's X's perception of agent Y's benevolence towards it, partly a function of agent X's perception of agent Y's capability toward task T, partly due to balance of tit-for-tat [9] and partly due to the power variations with others. This approach to conceptualizing trust lends itself to formulating delegation between two individuals, which requires trust between delegator and delegee [4, 11].

An agent's autonomy towards a task is affected by its capability and the sense of freedom it receives from other agents [7, 10]. This sense of freedom can be approximated by a combination of factors such as power and trust it receives from others. Personal autonomy of agent is considered to be exogenous sources, which might be powers, the agent perceives from other agents that are at different power levels and the permissions it gets from other agents. Generally, autonomy can interpreted as a combination of socially warranted freedoms and internally perceived freedoms. When individuals are in groups, organizations, and institutions, their individual autonomy is altered by the properties of the group. Autonomy is subject to constraints and context of the agent's environment and as such is not strictly determined by interpersonal dependencies.

In the remainder of this paper we will begin by elaborating our model of trust, autonomy, and delegation. In section three, we discuss our implementation we have used for our experimental results. In section four, we describe a series of experiments that illustrate our models. In section five, we draw some conclusions.

2. A MODEL OF TRUST, AUTONOMY, AND DELAGATION

Our model of trust is aimed at capturing a precondition to the formation of intentions to delegate a task, i.e., asking for a task to be done by another agent. An agent's assessment prior to delegation may include an analysis of risk and utilities, creating an intermediate notion of trusting value, prior to adoption of an intention. In most applications, trust has the consequence of reducing the need for the trusting agent to supervise or monitor the trusted agent.

The variety of definitions has added to the confusion about, and misconceptions of trust. In multi-agent systems, trust has been related to models of other social notions such as autonomy, delegation, dependence, control, and power, which influence interactions between agents. In this paper, we treat trust as a dyadic relation, i.e., the amount of trust each agent has on other agents. We define Trusting value to be the amount of trust an agent has on other agents with respect to a particular task [8]. This value among the agents is calculated by the following expression:

$$\boxed{\text{Trusting value } (A, B, t) = (1/3)*[\text{capability}(B,t) + \text{benevolence}(B, A, t) + 10* DH(A, B)]} \quad (1)$$

Here A, B are agents and t is the task to be performed by agent B. capability(B, t) is the agent B's ability to perform a task t and we assume both A and B perceive the same value. benevolence(B, A, t) is agent B's (i.e. trustee's) level of well wishing towards agent A (i.e. trusted) in performing a task t. DH(A, B) is the number of times agents A and B have agreed to the delegation request from one another after internally weighting all other considerations by the sum of number of times agents A and have made a delegation request. The DH value range from 0.0 to 1.0. DH is only computed when agents interact at the interpersonal level [9].

Autonomy value for an agent is the amount of trust the agent has for itself to perform a task and is computed by the following expression.

$$\boxed{\text{Autonomy value } (A, t) = (1/3)*[\text{capability}(A, t) + \text{Average } (T) + 1/(n-1)* \text{ Balance of reciprocity}]} \quad (2)$$

Balance of reciprocity for an agent A is counting two values and subtracting two values:

- Add the number of times delegated tasks by agent A has been agreed upon divided by the number of such agents
- Add the number of times agent A has made a delegation request regardless of accepting that request divided by the number of such agents
- Subtract the number of times agent A has agreed to a delegation request by another agent divided by the number of such agents making the request
- Subtract the number of times agent A has been asked for delegation regardless of whether A has agreed to work on the task divided by the number of such agents.

capability(A, t) is the agent A's ability to perform a task t. Average(T) is the average trust of all the agents on agent A and is measured by

$$1/(n-1) \sum_{i=1}^{n} T_i$$

where $T_1, T_2 \ldots T_n$ are the trusting values of the agents on agent A on a particular task t. The amount of trust an agent has on itself determines its competence for performing a task. We call this *autonomy* value of the agent as trusting value of an agent on itself. Equation 2 of an agent is same as the trusting value of the self-agent. Obviously, Equation (1) affects Equation (2). Agents use the trusting values to evaluate their own autonomy in deciding and performing the tasks. Power affects both trust and autonomy indirectly.

Autonomy is compared with the trusting values of all the agents to determine which agent should perform a task. An agent selects a task for which its autonomy is highest from the overall tasks determined. The agent then selects for which other agent it should perform the task depending on the highest power among the agents that assigned the task. When the powers are similar then the agent randomly selects an agent for which it performs the task. Every agent has an individual task assigned to perform. The autonomy of an agent to perform the pre-defined task is compared with the value of autonomy of the agent for whom it selected to perform the task. If the agent's self-autonomy is higher than the autonomy of the selected agent then it performs its own task for itself. When multiple agents determine to perform a unique task, an agent with higher power is performs the task. For agents with equal powers their autonomies with respect to the task are compared. For agents with equal autonomy their capabilities with respect to the task are compared and the agent with the higher capability performs the task. If the agent's capabilities are equal, the task is performed by one of the agents selected randomly.

The following sub-sections discuss the power homogeneity and commitment among agents.

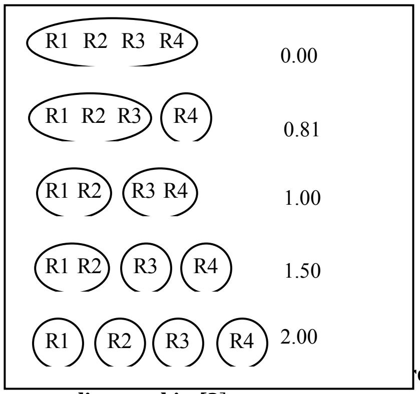
2.1. Power homogeneity

Our goal is to devise a metric that differentiates a group with agents at different power levels. Tucker Balch's social entropy differentiates the agents with different group size however; it does not include power [2]. We extend the principle of social entropy to differentiate among groups to different power profiles. For simplicity, the power homogeneity is represented as PH and is computed by the following expression.

$$PH = \left| \sum_{i=1}^{m} (p_i + w_i) \log_2 (p_i + w_i) \right| \quad (3)$$

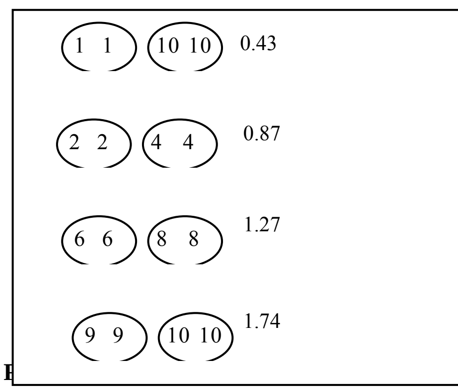
where m is the number of groups. p_i is defined as ratio of number of agents in a group i over the total number of agents. w_i is defined as the ratio of power level of the group over the maximum power level of all the groups.

The following clarifies the power homogeneity. Figure 1 shows the social entropy of different groups where there are 4 agents. In power homogeneity no two groups share the same agent.



roups
as discussed in [2]

Figure 2 shows the power homogeneity of agents having social entropy of 1.00 but different power levels. There are two groups with two agents in each group. The numbers in ovals show their power levels.



F **wer**
levels with same entropy

There are two special cases to be considered for power homogeneity: when there is a single group with all the agents having the same powers and when all the agents are at different levels.

Case 1: When there is a single group with all the agents having the same power the social entropy of these agents will be 0.00 and the power homogeneity for such groups will be the same for the agents with different power levels. In order to differentiate among such groups at different power levels we set the power homogeneity (PH) to the power level of the group.

Case 2: When power of agents are different from one another. In such case it would be easy to contrast the power homogeneity of agents when the organization is very small. As the number of individual in the group increases comparison is harder. In the rest of this paper we will overlook such groups.

An example is considered to illustrate the power homogeneity expression among agents with different power levels but at the same social entropy level. Let there be four agents A ={A1, A2, A3, A4} forming two groups G1 and G2 with two agents in each group. The social entropy is 1.00. Let the power level of G1 be 1 and power level of G2 be 10 respectively. The power homogeneity as calculated from equation 3 is:

i) Power of G1 = 1, Power of G2 = 10

P1 = 2/4 = 0.5, w1 = 1/10 = 0.1

P2 = 2/4 = 0.5, w2 = 10/10 = 1

$$PH(A) = +/-\sum_{i=1}^{2} (p_i + w_i)\log_2(p_i + w_i)$$

$$= +/-(((p_1 + w_1)\log_2(p_1 + w_1)) + ((p_2 + w_2)\log_2(p_2 + w_2)))$$

$$= +/-(((0.5 + 0.1)\log_2(0.5 + 0.1)) + ((0.5 + 1)\log_2(0.5 + 1)))$$

$$= 0.43$$

ii) Power of G1 = 2, power of G2 = 4

P1 = 2/4 = 0.5, w1 = 2/4 = 0.5

P2 = 2/4 = 0.5, w2 = 4/4 = 1

$$PH(A) = +/-\sum_{i=1}^{2} (p_i + w_i)\log_2(p_i + w_i)$$

$$= +/-(((p_1 + w_1)\log_2(p_1 + w_1)) + ((p_2 + w_2)\log_2(p_2 + w_2)))$$

$$= +/-(((0.5 + 0.5)\log_2(0.5 + 0.5)) + ((0.5 + 1)\log_2(0.5 + 1)))$$

$$= 0.87$$

iii) Power of G1 = 9, power of G2 = 10

P1 = 2/4 = 0.5, w1 = 9/10 = 0.9

P2 = 2/4 = 0.5, w2 = 10/10 = 1

$$PH(A) = +/-\sum_{i=1}^{2} (p_i + w_i)\log_2(p_i + w_i)$$

$$= +/-(((p_1 + w_1)\log_2(p_1 + w_1)) + ((p_2 + w_2)\log_2(p_2 + w_2)))$$

$$= +/-(((0.5 + 0.9) \log_2 (0.5 + 0.9)) + ((0.5 +1) \log_2 (0.5 + 1)))$$

$$= 1.55$$

As shown above, the PH values increase with the increase in the power levels and decrease with decrease in the power levels of the groups. The values of the PH do not have a range as it varies with varying number of agents.

2.2. Commitments

An agent selects a task and decides to perform that task for another agent taking power, autonomy and capabilities into consideration. Since a combination of factors is considered, an agent's decision might not always coincide with capability or trust. When exchange between agents coincides with trust and does not make sense in terms of capabilities then we say it is a case of *commitment*. Giving trust primacy over capability is commitment.

When the difference in capabilities between agents in a delegation exchange is greater than the difference in their trust then the exchange is said to be between trusted partners, which is motivated by commitment, otherwise it is considered to be a non-commitment or missed commitment. Let C represent the capability, T represent the trusting value, and Δ stands for difference in values.

$$\Delta T > \Delta C \rightarrow \text{Commitment}$$

$$\Delta C > \Delta T \rightarrow \text{non-Commitment}$$

where ΔC represent difference in the capabilities and ΔT represent difference in the trust.

Commitments among the agents are proportionally related to agent autonomy levels. Also commitments are proportionally related to power homogeneity. Experimental results in Section 4 support these two hypotheses.

```
1.Initialize the values of capability matrix (C[][] ) to random
  values between 0 to 10.
2.Initialize the values of Benevolence (B[][][] ) to 0.0 initially.
3.while (tasks remain) {/*main body of the algorithm*/
4.  for all agents and tasks { /* trusting values */
5.    if (a = b) /* a, b – variables stand for agents*/
          TV[t][a][b] = (1/3)*[C[a][t] + average(T) +
                  (Balance of reciprocity)/(n-1)]
6.    else TV [t][a][b] = (1/3)*[ C[a][t] + B[t][a][b] + 10*DH(a,b)]
7.    A[a][t] = (1/3)*[C[a][t] + average(T) +
        (Balance of reciprocity)/(n-1)] /* autonomy */
8.   compare A[][] with TV[][][] to find the suitable agents
     performing task t
9.   perform task t for the agent that has the highest power
     after comparing power of agents that assigned task t.
10.   compute the number of tasks being executed per
      iteration and unsuccessful attempts.
11.   C[][] = C[][] + i   /*Update C[][]  with success */
      C[][] = C[][] – i  /*Update C[][] with failure */
12.   B[][][] = B[][][] + i  /*Update B[][][] with success */
       B[][][] =  B[][][] – i   /*Update B[][][] with failure */
```

Figure 3. Algorithm to calculate the average trust and average autonomy

- AA[][] is an average autonomy of all the agents with respect to tasks and is used in plotting the graph.
- Average(T) is the average trusting values
- of all other agents with respect to self agent.
- TV[][][] is the matrix that holds the trusting values of agents with respect to tasks.
- B[][][] is the benevolence matrix of the agents.
- ATV[][][] is the average trusting values of all agents except the self-trusting values with respect to the tasks and is used in plotting the graph.
- B[][][] is zero initially but varies with time.
- DH is the delegation of harmony among the agents.
- Balance of reciprocity is defined in the equation 2.
- n is the number of agents

3. SIMULATION

In our implementation simulation, N agents considered N tasks repeatedly, i.e. each agent has its own task, which is same in each time period. This does not mean that each agent has to perform the assigned task. Agents may perform tasks assigned to other agents. The tasks are performed in part based on the power of the agents that assigned the tasks and the power of the agent who is considering the task. The power ranges from 0 to 10 that represent the power of the agents in this network. The power of the agents is fixed. The ranges of capability and trust are between 0.0 and 10.0. In our simulation we assume agents in general perform certain tasks and develop trust, capability, and benevolence in performing. In the algorithm, we focus on how the power of an agent affects the trust and autonomy and the commitments among them in performing the tasks. The pseudo code of our simulation is shown in Figure 3.

The success or failure of an agent can be determined by comparing the capability values of an agent with a randomly generated number ranging between 0 to10. If the random number has a value greater than the capability value of an agent, it is considered as a failure and if the number is smaller, then it is considered as a success. An agent may perform one task each time and no two agents can do the same task. The capability, trust and relations among the agents are updated with the success or failure in performing the tasks. Before the values are updated the average autonomy and trusting values of the agents are calculated to observe a relation between the two with respect to the power among the agents.

4. EXPERIMENTS AND SICUSSIONS

This section presents results of using our abstract simulation of agents and tasks. Two different experiments were performed. In the first experiment the results were observed for 25 units of time. In each time unit the average autonomy and average-trusting values were noted. Autonomy and trust of agents with three different power levels in the same group are considered and compared. The average trust and average autonomy of the agents were computed with group of agents that have equally low powers among them, group of agents with equally medium powers in second and agents with equally high powers in the third group in the network. As the range of the power is set from 0 to 10, a value of 1 represents the lowest power, a value of 5 represents medium power and a value of 10 represents highest power. These powers are fixed among the agents in the groups over 25 simulation units. Figure 4 shows the results of different power levels in the same group that effect the agent's autonomy and trust.

From Figure 4 we observe that the average trust and average autonomy values among the agents with equally medium powers in the second group were lower than the first and third groups where power level among agents are lowest and highest. This is because, the agents in the group of medium power level tried to perform and assign most of the tasks to the agents of their own power level. The group of agents with low power level tried to perform more tasks for other agents and that improved their ability in performing the tasks for other agents thereby increasing their autonomy. As the agents in a group are improving in their autonomy levels the trust among the agents within the group increases. The agents in the group of high power levels improved their autonomy and trust levels by convincing other power level agents to perform the tasks they assigned to them. Figures 5 and 6 support our discussion. The fluctuations in the curve may also be due to variations in capability of the agents. Benevolence value started at zero but changed over time with the delegation of tasks by the agents.

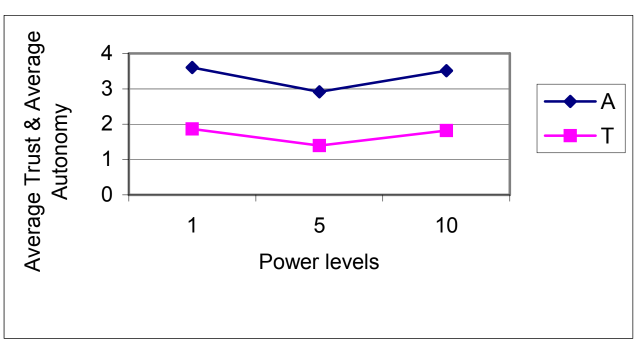


Figure 4. Average trust and Average autonomy of agents at different power levels

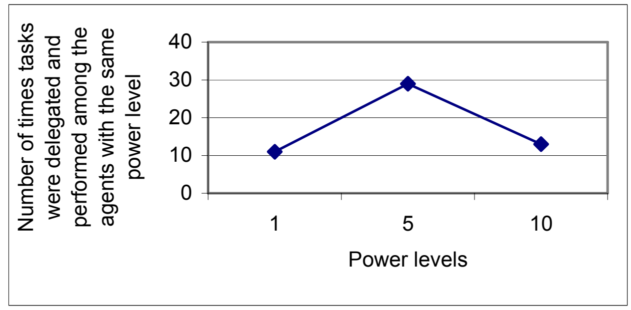


Figure 5. Number of tasks were delegated by and performed by agents in the same power level

Figure 5 shows the number of times an agent performed tasks for other agents in the same power level over 25 units of time. We observe from figure 5 that agents in the medium power level group tried to perform the most tasks for the agents with their own power level while the other power group agents have performed fewer tasks determined by themselves. Figure 6 shows the number of times the agents assigned tasks for agents of same power level group over 25 units of time. From Figure 6 we observe that agents in low and high power level groups assigned the tasks moderately fewer number of times to the agents that have the same power than the agents with medium power levels.

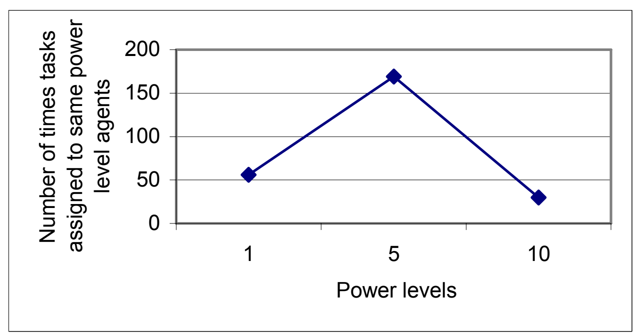


Figure 6. Number of tasks assigned for agents in the same power level group

We say that autonomy and trust of agents that perform and assign tasks to the other agents with different power levels will improve faster as they gain the ability to perform different tasks for different power level agents and the ability to convince other agents to accept the assigned tasks.

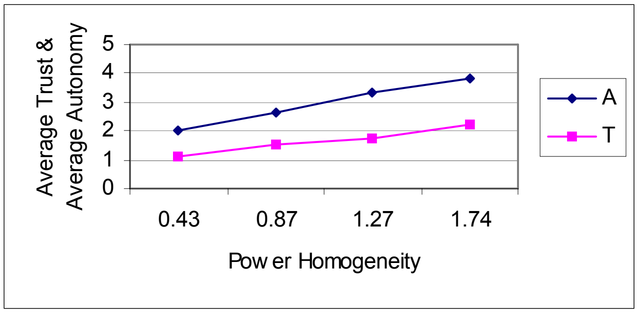


Figure 7. Average trust and Average autonomy of agents with power homogeneity

In the second set of experiments average trust and average autonomy of the agents with different power homogeneity were observed. The results were observed at 25th time unit. Figure 7 shows how the average trust and average autonomy of the agents vary with different power homogeneity with the same social entropy. The experiments were performed with 4 agents where there are two groups and each group has two agents. The social entropy for this set is 1.00. The two groups have different power levels. It is observed from figure 7 that the average trust (T as shown in figure 7) and average autonomy (A as shown in figure 7) of the agents with different power homogeneity increases with increase in the power homogeneity. This is because as the level of power increases, agents becomes more autonomous in performing or getting tasks done by other agents and when the difference in the power levels of the groups is less the agents would be more feasible in understanding one another which increases their trust level when working with agents that have less difference in their power levels. The same principle holds for any group levels with same social entropy and different power homogeneity.

In the third set of experiments two tests were performed to observe how commitment and autonomy varies among the group of agents with different power homogeneity (PH). The experiments were performed with 4 agents in the group and the results were observed at 25th time unit. Figure 8 shows the results of a test that computed the level of commitment among the agents at different power equals. It is observed from Figure 8 that the level of commitment increases with increase in the power homogeneity. There was a gradual increase in the commitment curve. The level of commitment was 42 when power homogeneity was 0.43. When the power homogeneity was 4.02 the level of commitment was 55 and the commitment was 64 when the power homogeneity of the group of agents is 4.36.

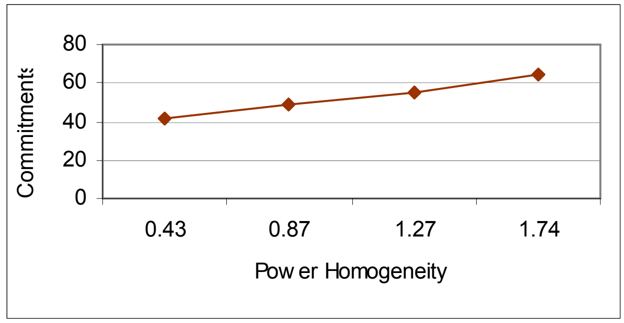


Figure 8. Commitments at different power Homogeneity

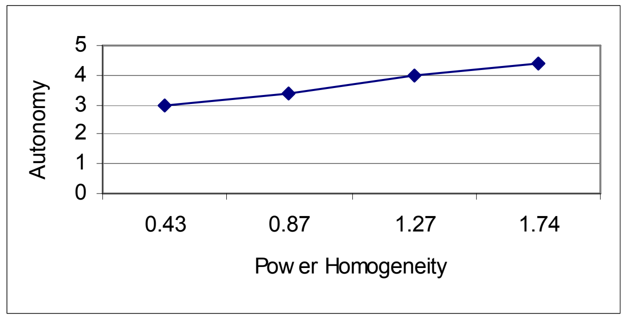


Figure 9. Autonomy at different power Homogeneity

Another test was performed to observe the relationship between autonomy and commitment. Figure 9 shows average autonomy at different power homogeneity levels. The average autonomies of a group of agents were observed simultaneously with the commitments at different power homogeneity levels. It is observed that the average autonomy increases gradually with increase in power homogeneity among agents. The average autonomy at PH=0.43 was 2.95. At PH=0.87 the average autonomy was 3.41 and at PH=1.74 the average autonomy was 4.36 respectively. Changes in autonomy increase predict changes in commitment among the agents at different homogeneity levels. In other words, autonomy is a predicator of commitments.

6. CONCLUSIONS

A simple model of autonomy and trust was presented and shown to be subject to power levels among agents. Presenting many parameters will have obscured our observations. This model is deliberately kept simple to illustrate the role of relative power between autonomy and trust. An agent experiences autonomy with respect to a task if it is capable of performing or is capable of convincing other agents to perform tasks. The notion of power homogeneity is sown to affect trust and autonomy.

We have seen from our experiments that when power levels are high and low agents tend to have higher trust and autonomy levels than the agents at medium power levels. Also, when the power homogeneity is high the trust and autonomy levels are high. It is observed that commitments mirrored changes in autonomy in groups of different power homogeneity.

6. ACKNOWLEDEMENTS

This work is supported by AFOSR grant F49620-00-1-0302.

7. REFERENCES

[1] Abdul-Rahman A, Hailes S. Supporting Trust in Virtual Communities. In: Proceedings *Hawaii International Conference on System* Sciences *33, Maui, Hawaii, 4-7 January-2000*.

[2] Balch T. Social Entropy: a New Metric for Learning Multi-robot Teams. In: Proceedings *Flairs-97.*

[3] Brainov S, Sandholm T. Power, Dependence and Stability in Multiagent Plans. American National Conference on Artificial Intelligence (AAAI'99), Orlando, AAAI Press, pp. 11-16, 1999.

[4] Castelfranchi C & Falcone R. Principles of trust for MAS: Cognitive anatomy, social importance, and quantification. In: Proceedings of *the Third International Conference on Multi-Agent Systems,* Paris, France, pp 72-79, 1998.

[5] Hardy C, Phillips N and Lawrence T. Distinguishing trust and power in Interorganizational relations: Forms and Facades of Trust. In: *Trust within and between organizations,* Christel Lane and Reinhard Bachmann *(eds)*, Oxford University press, pp 64-87, 1998.

[6] Hexmoor H. A Model of Absolute Autonomy and Power: Toward Group Effects, In Journal of Connection Science, Volume 14, No. 4. Taylor & Francis Ltd, 2003.

[7] Hexmoor H. Stages of Autonomy Determination. In: *IEEE Transactions on Man, Machine & cybernetics-Part C* (SMC-C), Vol. 31, No. 4, pp 509-517, November-2001.

[8] Hexmoor H and Poli P. Benevolence, Trust and Autonomy. In: Proceedings of the International Conference on Artificial Intelligence (IC-AI'03) Las Vegas, Nevada, 2003.

[9] Hexmoor H and Poli P. Effects of Reciprocal Social Exchanges on Trust and Autonomy. In: Proceedings of Third International Conference on Intelligent Systems Design and Applications (ISDA-03) Tulsa, Oklahoma, 2003.

[10] Hexmoor H and Vaughn J. Computational Adjustable Autonomy for NASA Personal Satellite Assistants. In: *ACM SAC-02*, Madrid, 2002.

[11] Sichman J S, Demazeau Y, and Conte R, Castelfranchi C. A Social Reasoning Mechanism On Dependence Networks. In: Proceedings of *the Eleventh European Conference on Artificial Intelligence,* 1993.