

Università degli Studi di Salerno

FACOLTÀ DI SCIENZE MATEMATICHE FISICHE E NATURALI
Corso di Laurea Triennale in Informatica



Tesi di Laurea in Informatica

Social Network Advertising: Un algoritmo Greedy sulle Clique Pesate

Candidato:

Jacek Filipczuk

Mat.: 0512100361

Relatori:

Dott. Francesco Carrabs

Anno Accademico 2011-2012

Ringraziamenti

A mia madre, perno fermo e inamovibile della mia vita, sempre presente sia nei momenti di bisogno sia in quelli in cui non dovrebbe mettere il naso. Lei, instancabile nei suoi consigli quanto nella fiducia in me riposta. Lei, l'unica donna che senza sosta cerca tutti i giorni di mettere un po' di sale nella mia zucca.

A Emanuele, Infimo per gli amici, amico fedele e insostituibile.

Al SaintHappy che con la sua cupola di protezione regala saggezza e sicurezza a tutti noi.

A Andrea, Paolo, Gianluca, Davide, Alessandra, Giovanni, Demia e tutti gli amici che mi hanno accompagnato durante questo percorso.

Al Dott. Francesco Carrabs che ha avuto fiducia in me e nelle mie capacità fin dall'inizio.

A loro e a tante piccole cose come la musica che mi hanno dato la forza e la determinazione di continuare.

A tutti voi....

Grazie Di Cuore

Indice

1. Introduzione.....	4
2. Social Network Analysis.....	7
2.1 Storia dei Social Network.....	9
2.2 Impacts of Ads in Social Networking Sites.....	11
2.3 Combining Behavioral and Social Network Data for Online Advertising.....	16
2.4 Avoidance of Advertising in Social Networking Sites.....	18
3. The Maximum Clique Problem.....	20
4. Definizioni e Modello Matematico.....	24
4.1 Applicazioni Pratiche.....	25
5. Euristiche e Pseudocodice.....	30
5.1 Euristiche Smart Clique.....	31
5.2 Multistart.....	31
5.3 Dynamic Local Search.....	31
6. Testing.....	38
7. Considerazioni Finali e Sviluppi Futuri.....	40
8. Bibliografia.....	42
9. Tabella Risultati Testing.....	47

1. Introduzione

Nel campo dell'informatica spesso bisogna risolvere numerosi problemi di elevata complessità. Uno di questi è il Problema della Massima Clique di un grafo. Un grafo $G=(V,E)$ non è altro che un insieme di vertici e archi che collegano questi vertici. Una clique invece è un sottoinsieme dei vertici di G , ovvero $V' \subseteq V$, tale che esiste un arco che collega ogni coppia di vertici appartenenti a essa. La clique è uno dei problemi più studiati nel campo della matematica perché la sua risoluzione offre sempre numerose informazioni interessanti. Al problema della Clique è associato il corrispondente problema di decisione. Il problema di decisione consiste nel determinare se un grafo G contiene una clique di taglia k , ovvero con k vertici. Questo problema risulta difficile ed è stato dimostrato essere un problema NP-Completo e di conseguenza intrattabile.

Il problema della Massima Clique è uno dei problemi più studiati nel campo della Teoria Computazione e consiste nel determinare la clique di taglia massima di un grafo.

Esistono inoltre numerose altre varianti del problema della clique, come ad esempio:

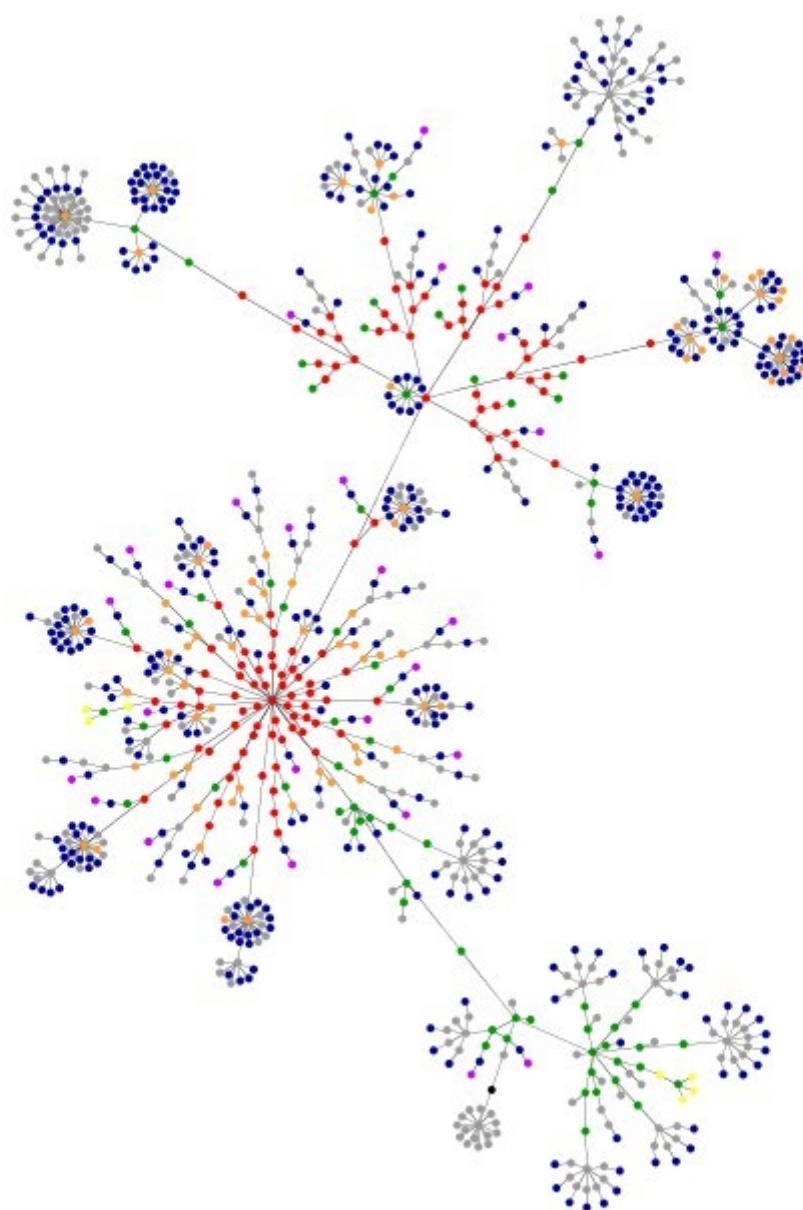
- Trovare la clique di peso massimo in un grafo pesato
- Elencare tutte le clique massimali di un grafo
- Risolvere il problema di decisione della Clique
- Elencare tutte le clique massime di un grafo

Questo lavoro di tesi introduce un nuovo approccio per risolvere il problema della clique di massimo peso. Più precisamente è stato sviluppato un nuovo algoritmo greedy per il Problema della Massima Clique Pesata.

Per prima cosa viene presentato il campo di studio in cui può essere maggiormente applicato l'algoritmo sviluppato, ovvero L'analisi dei Social Network. Viene data al lettore una panoramica generale del Social Network Analysis, descrivendo la sua nascita, lo sviluppo e soprattutto gli oggetti di studio e le sue possibili applicazioni. Inoltre si è voluti porre una maggiore attenzione agli studi effettuati nelle reti sociali e all'uso di queste per la diffusione di annunci pubblicitari, oltre a tutte le motivazioni che stanno alla base delle scelte pubblicitarie effettuate dalla compagnie per massimizzare il proprio guadagno.

Successivamente vengono fornite al lettore le nozioni base necessarie per capire fin nei minimi dettagli il lavoro eseguito. Nel caso in cui il lettore abbia una discreta conoscenza della Teoria dei Grafi, allora può tranquillamente saltare la lettura del capitolo relativo e dedicarsi al resto del lavoro.

Di seguito viene mostrata una figura di esempio che rappresenta un tipico grafo. Ovviamente i grafi possono essere di tutte le dimensioni, alcuni con migliaia di nodi e archi.



In figura è mostrato un tipico esempio di grafo.

Nel capitolo seguente viene presentato il modello matematico sfruttato in questo lavoro. Vengono descritti nei dettagli tutti i vincoli sfruttati e le caratteristiche del modello necessari allo studio del problema. L'analisi del modello viene seguita dalla sezione in cui vengono mostrati vari esempi di applicazioni pratiche dell'algoritmo sviluppato.

Il capitolo sull'euristica risulta essere il fulcro del lavoro di questa tesi. In esso vengono descritte le caratteristiche dell'euristica che sta alla base del lavoro. Inoltre

vengono mostrati i dettagli dell'algoritmo sviluppato insieme al pseudocodice, tutto per presentare al lettore una visuale chiara e precisa del lavoro svolto.

Il penultimo capitolo consiste nella descrizione dei risultati ottenuti eseguendo l'algoritmo sviluppato su istanze DIMACS, famose e sfruttate in tutto il mondo.

Infine nell'ultimo capitolo vengono descritte le conclusioni e una serie di possibili miglioramenti all'algoritmo proposto e alcuni possibili sviluppi futuri.

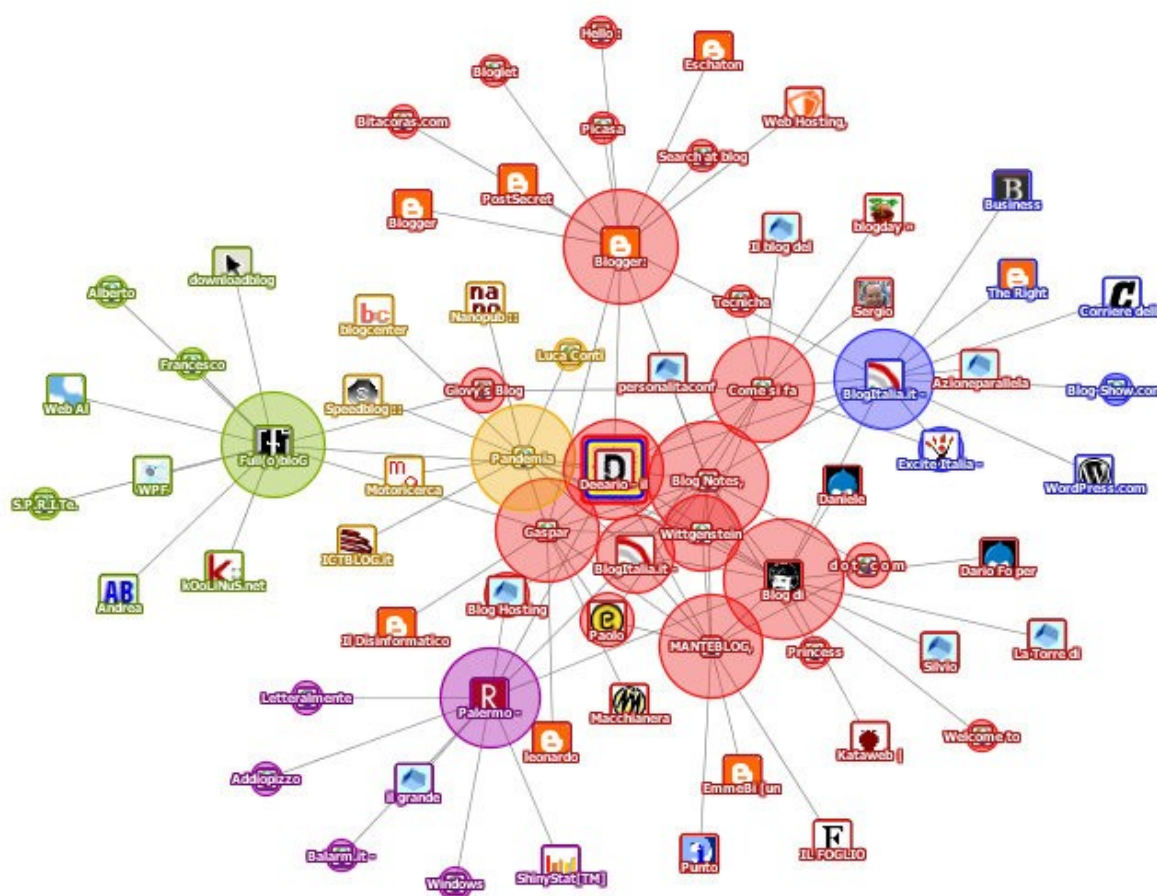
2. Social Network Analysis

La vita sociale è relazionale; ma solo quando, ad esempio, i bianchi e i neri usano particolari tipi di modelli comportamentali fra di loro, che la parola razza diventa una variabile importante.[Collins]

Le comunità scientifiche che studiano il comportamento e le relazioni sociali sono state molto attratte, negli ultimi anni, dalle nozioni di Social Network e Social Network Analysis. La maggior parte di questo interesse è dovuto alla concentrazione del Social Network Analysis sulle relazioni tra le entità e sui metodi e conseguenze di queste relazioni. Molti ricercatori hanno realizzato che la prospettiva dei network fornisce un approccio più preciso per la definizione di alcuni aspetti di ambienti sociali come quello politico o economico.

La differenza fondamentale tra una spiegazione basata sulle Social Network e una senza, consiste nella presenza di concetti e informazioni sulle relazioni tra gli individui.

L'esempio mostrato in figura rappresenta, tramite un grafo, i collegamenti esistenti tra pagine internet. Ovviamente l'esempio è molto ridotto, perché i possibili collegamenti sono numerosi ma grazie all'uso dei grafi, è possibile mettere in evidenza solo le informazioni richieste.



In figura è mostrato un piccolo grafo dei possibili collegamenti tra due pagine internet.

Lo studio dei Social Networks mette in risalto le teorie, i modelli e le applicazioni che sfruttano il concetto di relazione. Le relazioni definite tra le entità sono una parte fondamentale della Teoria dei Social Network. Oltre all'uso di concetti relazionali, è importante specificare gli altri concetti largamente sfruttati nelle Social Network :

- Gli Attori (così vengono chiamate le entità all'interno di una rete sociale) e le loro azioni sono viste come interdipendenti, di conseguenza non vengono mai considerati indipendenti o autonomi;
- I Legami relazionali tra gli attori sono considerati dei mezzi per il trasferimento di risorse (materiali o immateriali, come ad esempio le informazioni);
- I Modelli delle Reti che si concentrano sugli individui, visualizzano la struttura

- I Legami relazionali tra gli attori sono considerati dei mezzi per il trasferimento di risorse (materiali o immateriali, come ad esempio le informazioni);

- I Modelli delle Reti che si concentrano sugli individui, visualizzano la struttura

dell'ambiente della rete come se questa fornisse un'opportunità o una costrizione per l'azione individuale;

- I Modelli delle Reti concepiscono le strutture (sociali, economiche, politiche, ecc.) come un insieme delle relazioni esistenti tra gli attori.

La Social Network Analysis viene sfruttata molto nello sviluppo del modello, nella specifica o nel testing di un problema sulle reti sociali perché fornisce definizioni formali, misure e descrizioni utili a valutare modelli e teorie nelle quali il concetto principale è espresso come una relazione; fornisce anche un'analisi statistica di sistemi multirelazionali.

Alcuni degli argomenti su cui si concentrano gli studiosi delle Social Network sono:

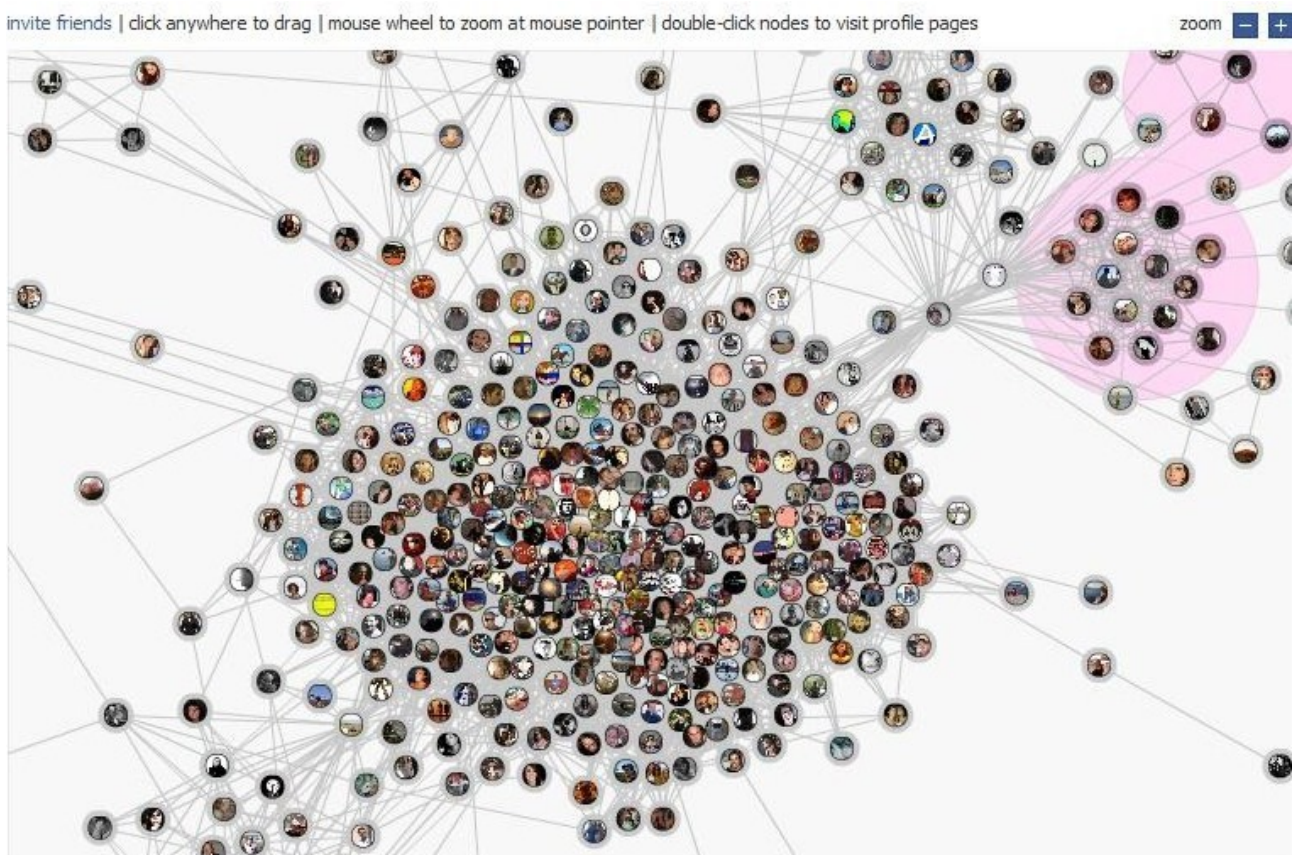
- Occupational Mobility (Breiger 1981c, 1990a);
- Impatto dell'urbanizzazione sui singoli individui (Fischer 1982);
- Il sistema politico ed economico mondiale (Snyder e Kick 1979, Nemeth e Smith 1985);
- Supporto Sociale (Gottlieb 1981, Lin, Woelfel e Light 1986, Kadushin 1966, Wellman, Carrington, e Hall 1988, Wellman e Wortley 1990);
- Diffusione e adozione delle Innovazioni (Coleman, Katz, e Menzel 1957–1966, Rogers 1979) ;
- Cognizione e percezione sociale (Krackhardt 1987a, Freeman, Romney e Freeman 1987) ;
- Approvazione e Influenza Sociale (Friedkin 1986, Friedkin e Cook 1990, Doreian 1981, Marsden 1990a) .

L'approccio usato dai Network differisce in maniera fondamentale dagli standard usati nella ricerca sociale e in quella comportamentale. Invece di focalizzarsi sugli attributi delle singole unità, sulle associazioni tra questi attributi o sulla loro utilità nel predire informazioni riguardanti altri attributi, l'approccio usato dai Social

Network si concentra sulle caratteristiche della struttura relazionale presente tra le unità in esame. L'obiettivo è quello di capire come le proprietà strutturali di un ambiente sociale (politico o economico) influenzano le caratteristiche osservate o le associazioni tra queste caratteristiche.

Dato un insieme di attori, la Social Network Analysis può essere usata per studiare la struttura delle variabili misurate tra gli attori. La struttura relazionale di un gruppo o di un sistema sociale più grande consiste nell'insieme dei modelli relazionali esistenti tra gli attori dell'insieme. Difatti la parola stessa Social Network si riferisce all'insieme degli attori e alle relazioni fra di loro. Un analista dei Network ricerca il modello di queste relazioni per poter rappresentare la struttura del gruppo. In questo modo è possibile sfruttare le informazioni riguardanti la struttura per capire che impatto essa può avere sui singoli individui presenti all'interno del gruppo.

2.1 Storia dei Social Network



In figura un esempio di grafo delle amicizie di Facebook.

La Social Network Analysis è, senza dubbio, una materia interdisciplinare. I concetti della Social Network Analysis sono stati sviluppati grazie all'unione della Social Theory con rigorose metodologie matematiche, statistiche e computazionali. Come Freeman (1984), Marsden e Laumann (1984) hanno documentato, sia le scienze sociali che quelle matematiche e statistiche hanno tratto notevoli vantaggi grazie agli sforzi di ricercatori di queste discipline che collaboravano tra di loro.

Successivamente i concetti principali di relazione, network e struttura si sono sviluppati quasi autonomamente in numerose scienze sociali e comportamentali. I fondatori della Social Network Analysis provengono in ogni caso dall'ambito della Sociologia e della Psicologia Sociale. Infatti molte persone attribuiscono il primo utilizzo della parola Social Network a Barnes (1954). La nozione di una rete di relazioni che collega entità sociali ha trovato numerosi utilizzi nelle scienze sociali. Inoltre numerosi concetti appartenenti alle Social Network Analysis si sono sviluppati grazie a ricercatori che hanno dovuto usarli per risolvere svariati problemi empirici o teorici all'interno dell'ambito scientifico. Il fatto che così tanti scienziati provenienti da discipline così differenti hanno scoperto, quasi simultaneamente, la prospettiva delle Social Network, non deve sorprendere il lettore. La Social Network Analysis fornisce risposta a un gran numero di problemi, spaziando nell'ambito di numerose diverse discipline.

Alcuni scienziati hanno interpretato la Social Network Analysis come una serie di procedure analitiche che sembrano quasi separate completamente dai principali concetti teorici ed empirici. Alcuni metodi particolari possono anche sembrare privi di fondamenti teorici a causa del fatto che è possibile sfruttarli in un vasto gruppo di problemi provenienti da diverse discipline. Questa concezione è falsa, perché la Social Network Analysis è fondata su importanti fenomeni sociali e concetti teorici.

La Social Network Analysis fornisce anche significati concettuali per poter pensare in modo formale al mondo sociale. Freeman ha affermato che i metodi della Social

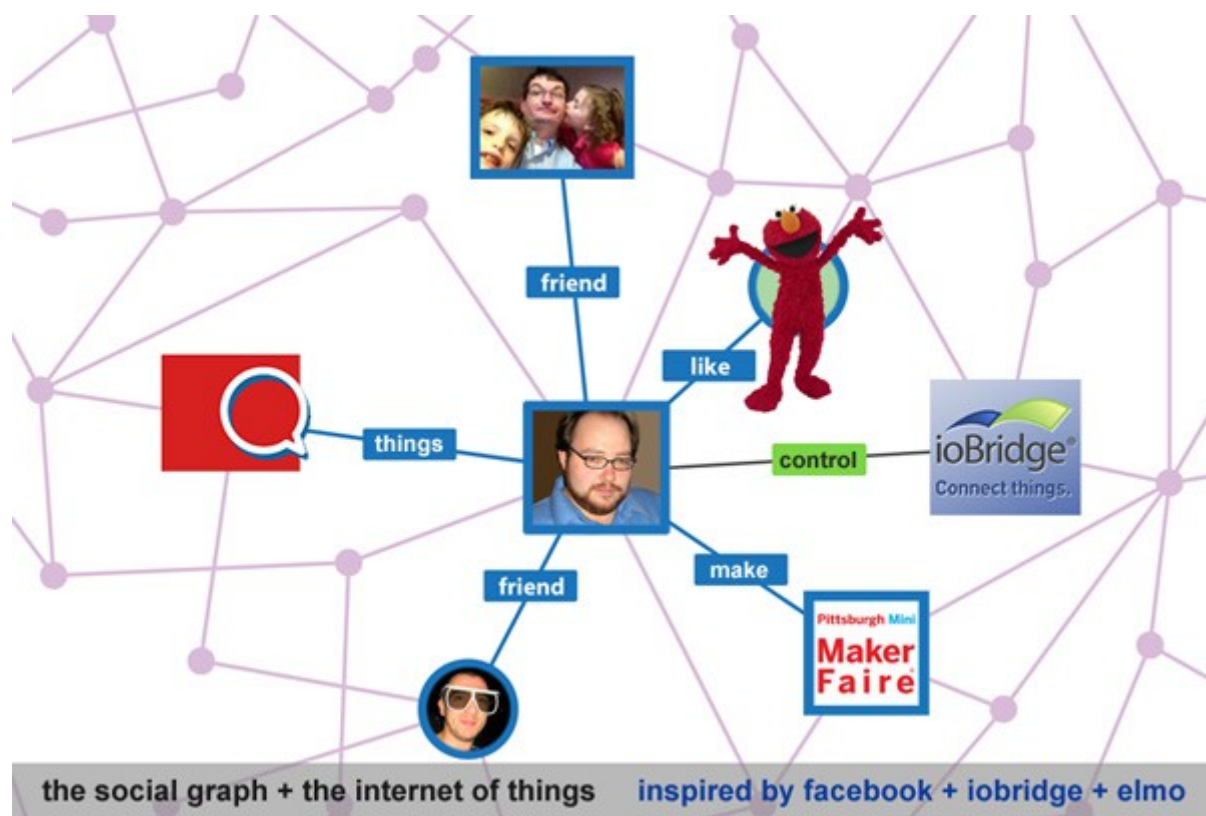
Network Analysis forniscono una dichiarazione formale delle proprietà e dei processi sociali. Successivamente queste dichiarazioni devono essere definite in modo preciso e concreto. Una volta che queste dichiarazioni, o concetti, vengono definiti con precisione, è possibile discutere in modo scientifico e rigoroso sul mondo sociale. Freeman ha affermato anche che il “gruppo” e il “ruolo sociale” sono due idee centrali che servono soltanto a sensibilizzare i concetti utilizzati, a meno che non vengono prima definiti in modo formale nei termini delle Social Network.

La nascita della Social Network Analysis viene però attribuita a Moreno che nel 1953 ha inventato il Sociogramma. Questa innovazione ha segnato la nascita della Sociometria (il precursore della Social Network Analysis e della Psicologia Sociale). Guidato dall'interesse nel capire i comportamenti psicologici e sociali degli esseri umani, Moreno fu spinto a inventare un nuovo metodo per rappresentare la struttura interpersonale dei gruppi di persone: il Sociogramma. Il Sociogramma è una rappresentazione grafica in cui le persone (o più generalmente le entità) sono rappresentate come punti in uno spazio bidimensionale e le relazioni tra le coppie di punti sono rappresentate tramite archi non direzionati tra i punti stessi. Moreno ha affermato, nel 1953, che prima dell'avvento della Sociometria nessuno sapeva come poteva essere rappresentata una struttura interpersonale di un gruppo. Moreno non si fermò al Sociogramma, la sua necessità di modellare importanti fenomeni sociologici lo portò a formulare due dei concetti fondamentali della Social Network Analysis: una rappresentazione grafica della struttura di un gruppo e un modello statistico dei risultati strutturali. La rappresentazione grafica comprende il Sociogramma e rappresentazioni a due o più dimensioni che vengono tutt'ora sfruttate nella Social Network Analysis. La scoperta del Sociogramma come strumento usato per lo studio delle strutture sociali ha portato a una rapida introduzione di tecniche analitiche. Nello stesso periodo ci fu la scoperta dell'utilizzo di matrici per la rappresentazione dei dati delle social network. Queste scoperte hanno aperto la strada ai matematici nello studio dei sistemi sociali i quali successivamente hanno sviluppato concetti importanti come la centralità degli attori o dei gruppi. I tre principali concetti

matematici sfruttati nei metodi per le reti sono la teoria dei grafi, la teoria della probabilità e statistica e i modelli algebrici. La teoria dei grafi, in particolare, fornisce sia un'adeguata rappresentazione che un insieme di concetti che possono essere sfruttati per lo studio delle proprietà delle reti sociali. La teoria della Statistica è diventata importante nel momento in cui sono cominciati gli studi sulla reciprocità, sulla mutualità, sull'equilibrio e sulla transitività presenti tra gli attori di un gruppo. La maggior parte della concentrazione si sviluppò sullo sviluppo di modelli matematici atti a descrivere il comportamento di triple di attori. Infine i modelli algebrici sono largamente utilizzati nello studio di reti multirelazionali. Questi modelli sfruttano le operazioni algebriche per studiare le combinazioni di relazioni e fornire un ruolo alle reti strutturali. Le teorie che sfruttano le idee della Social Network Analysis si distinguono dalle altre perchè si concentrano sull'influenza che un attore di un gruppo genera sugli altri attori e non li considerano come singoli individui indipendenti dagli altri.

2.2 Impacts of Ads in Social Networking Sites

In figura è mostrato un esempio di tutte le informazioni che un Social Network raccoglie al suo interno.



I Social Networking Sites sono dei siti web che offrono la possibilità agli individui di creare un proprio profilo pubblico all'interno di un sistema, e di stabilire svariati tipi di connessioni e relazioni con altri profili presenti nel sistema stesso. Questo tipo di siti internet permette agli utenti di condividere i proprio interessi con gli altri, tramite la creazione di gruppi e forum, oltre a fornire numerosi altri servizi. La cosa che rende i Social Network unici non è la possibilità di incontrare nuove persone, ma piuttosto la capacità degli utenti di definire, articolare e visualizzare la propria rete sociale. Nella maggior parte dei Social Networking Sites le persone comunicano principalmente con altre persone che fanno già parte della loro rete sociale. I Social Networking Sites hanno sviluppato una notevole varietà di servizi da offrire agli utenti, anche se il servizio che funge da fondamento a questo tipo di siti internet è quello di visualizzare i profili degli utenti insieme alla loro lista degli amici, anche questi utenti del sistema. I profili sono delle pagine uniche in cui un utente è incoraggiato a descrivere se stesso, spesso con l'aiuto di brevi e semplici questionari. Le risposte contenenti informazioni sull'utente, saranno visualizzare in appositi campi sul profilo di quest'ultimo, insieme a foto o altro materiale multimediale che l'utente è

invitato a inserire.

La pubblica visualizzazione delle connessioni è un argomento fondamentale dei Social Network Sites. La lista degli amici di un utente serve a mantenere un collegamento con i profili di questi amici, e ogni altro utente in grado di visualizzare il profilo di un individuo, è abilitato anche a visualizzare la sua lista di amici, avendo così la possibilità di attraversare la rete sociale del proprietario del profilo. Alcuni siti offrono la possibilità di nascondere parti del profilo di un utente perfino a quei utenti abilitati di visitarlo, offrendo questo privilegio solo ad alcuni individui prescelti direttamente dal proprietario del profilo. Inoltre la maggior parte dei Social Network Sites offre servizi di messaggistica tra gli utenti e di condivisione di materiale multimediale.

Social Network Advertising è un termine che viene usato per descrivere una forma di pubblicità che si focalizza sui Social Network. Infatti i Social Network, essendo un punto di ritrovo per migliaia di utenti, diventano dei mezzi allettanti per le compagnie pubblicitarie.

Ci sono due modi per sfruttare i Social Network Sites da parte delle compagnie pubblicitarie. La pubblicità diretta, basata sulla rete delle amicizie, e la pubblicità indiretta, basata sulla creazione di “pagine” o “gruppi” (in altri termini profili pubblici) a cui gli utenti possono iscriversi.

Un Social Network Site offre all'utente la possibilità di rendere private alcune se non tutte le informazioni presenti sul proprio profilo, in generale i Social Network Sites cercano di offrire un servizio sempre più vasto e articolato per garantire all'utente il miglior modo di trovare nuove amicizie e di creare nuove relazioni o espandere quelle esistenti.

I Social Network Sites continuano a sviluppare nuovi servizi e a migliorare quelli esistenti, la creazione di un account però risulta gratuita, di conseguenza i Social Network Sites devono trovare un guadagno in altri modi. Il Social Network Advertising è il metodo più diffuso con il quale i Social Network Sites riescono a

guadagnare con il proprio lavoro. Infatti ci sono svariati modi in cui è possibile sfruttare i Social Network Sites per operazioni pubblicitarie, tutti che stimolano la creatività e la sperimentazione di nuove tecniche in grado di catturare l'attenzione dell'utente.

I Social Network Sites vengono, quindi, sfruttati per raggiungere tutti quegli utenti che evitano i messaggi pubblicitari trasmessi con altri mezzi di comunicazione come la televisione. Le compagnie pubblicitarie investono infatti numerose somme di denaro per poter utilizzare lo spazio pubblicitario fornito dai Social Network Sites.

In generale esistono quattro ragioni principali per cui i Social Network sono usati come mezzi pubblicitari.

1. Cost Efficiency: La pubblicità diffusa nei Social Network ha un costo di produzione relativamente basso, in confronto ai mezzi tradizionali. Inoltre le compagnie pubblicitarie hanno la possibilità di creare tecniche pubblicitarie molto creative e gratuite, come ad esempio la creazione di un profilo per un prodotto in un Social Network.
2. Large Reach: Siccome i Social Network hanno milioni di utenti attivi e sono sempre più popolari, diventano un ottimo mezzo per la diffusione pubblicitaria su larga scala. Il numero di utenti raggiunti da un Social Network è il più grande di qualsiasi altro mezzo di comunicazione di massa dello stesso tipo.
3. Targeted Advertising: Le compagnie pubblicitarie hanno accesso a una grande quantità di informazioni sugli utenti e quindi anche i loro interessi. Queste informazioni permettono alle compagnie di generare pubblicità mirate ai singoli utenti.
4. Time Spent Online: Gli utenti passano sempre più tempo su internet, specialmente sui Social Network, a spese di altri mezzi pubblicitari come la televisione e i giornali. Le compagnie pubblicitarie cercano, di

conseguenza, di raggiungere tutti quei utenti che passano il loro tempo su internet.

Negli anni passati molte persone hanno detto che i Social Network sono una perdita di tempo ma on il recente Boom le varie compagnie pubblicitarie hanno preso seriamente in considerazione i Social Network come nuovo mezzo di diffusione pubblicitario.

Infatti ecco alcune interessanti statistiche che dimostrano l'efficacia del Social Network Advertising:

1. Il 60% del traffico generato nei Social Network proviene da città medio piccole, anche se il traffico più alto viene generato da una metropoli: Mumbai;
2. Il maggior numero di utenti proviene da una fascia di età tra i 15-24 anni;
3. Il rapporto maschio-femmina degli utenti si aggira intorno a un 4 a 1 in favore del sesso maschile;
4. I media sociali in India raggiungono quasi il 60% dell'utenza online;
5. Facebook e Orkut contengono quasi il 90% degli utenti dei Social Network;
6. Più del 45% degli utenti accedono ai Social Network più di una volta al giorno. Gli utenti di Facebook, ad esempio, visitano il Social Network almeno 3 volte al giorno;
7. La maggior parte del tempo sui Social Network viene speso su giochi e applicazioni.

In questo modo le compagnie pubblicitarie hanno aumentato sia il loro guadagno sia quello dei Social Network che si sono ingranditi a loro volta.

Tuttavia, anche se i Social Network crescono ad un ritmo esorbitante, i Social Network Advertising non riescono a sfruttare tutto il loro potenziale. Infatti il Social

Network Advertising, per diventare un efficace strumento per la diffusione pubblicitaria, deve seguire alcuni criteri:

1. Age Group Filtering: Gli utenti durante la loro registrazione sono tenuti a fornire dettagli riguardanti la loro età. E' possibile sfruttare queste informazioni per classificare gli utenti dei Social Network in gruppi e, successivamente, indirizzare anche le migliori pubblicità in base allo specifico gruppo d'età.
2. Utilizzo di pagine di Celebrità: Un altro metodo per aumentare l'utenza raggiunta da un annuncio pubblicitario è quello di sfruttare le pagine di grandi celebrità. In questo modo l'annuncio viene visto da tutti i fan della celebrità e di conseguenza l'utenza raggiunta risulta maggiore.
3. Attirare gli utenti tramite scopi da raggiungere: Grazie alle nuove tecnologie i singoli utenti sono in grado di esprimere completamente la propria unica personalità sul proprio profilo online. Le compagnie pubblicitarie possono sfruttare questa voglia di espressione dell'utente per inserirlo nel processo pubblicitario all'interno dei Social Network. La Sony ad esempio ha incentivato gli utenti a giocare ad un gioco online in cui i protagonisti erano vampiri, tutto ciò per pubblicizzare un nuovo prodotto in uscita sul mercato. In questo modo gli utenti giocavano con la speranza di ricevere un piccolo premio da parte della Sony, ma nello stesso tempo pubblicizzavano il gioco ad altri amici e con esso il nuovo prodotto proposto dalla compagnia.
4. Pubblicità di Stagione: Sono pubblicità diffuse in particolari stagioni dell'anno in modo da massimizzare l'interesse dell'utente per il prodotto pubblicizzato. Ad esempio è più efficace pubblicizzare articoli da mare nel periodo pre-estivo piuttosto che in quello invernale. Si tende infatti a classificare le pubblicità in base alla domanda da parte degli utenti, la quale cambia a seconda della stagione dell'anno. In questo tipo di azioni pubblicitarie il tempismo è la chiave del successo.
5. Uso mirato di Calligrafie: La calligrafia presente all'interno di un annuncio

pubblicitario deve fare in modo che il messaggio arrivi dritto all'utenza interessata. Il colore e lo stile della calligrafia devono essere accattivanti e devono saper catturare l'attenzione dell'utente.

6. Dove e quando Pubblicizzare?: Gli annunci pubblicitari devono essere posizionati nel punto di una pagina dove c'è grande attività. Gli annunci devono capire gli interessi dell'utente e in che punti della pagina egli esegue le sue attività, e proprio in quei punti collocare le pubblicità.
7. Slot Temporali: Le attività di un utente cambiano nel corso della giornata, e con esse anche i suoi interessi. Gli annunci pubblicitari dovrebbero tenere conto anche della fascia oraria in cui essere visualizzati. Ad esempio una buona idea sarebbe quella di pubblicizzare posti di ristorazione nelle fasce orarie pre-pasto, oppure tenere conto delle fasce orarie in cui l'utenza attiva di un Social Network è maggiore.

Una compagnia pubblicitaria non riuscirà a guadagnare il necessario se non avrà pubblicizzato nel migliore dei modi il proprio prodotto. Seguendo i suggerimenti sopra proposti invece sarà in grado di massimizzare l'utenza raggiunta dai propri annunci pubblicitari.

2.3 Combining Behavioral and Social Network Data for Online Advertising

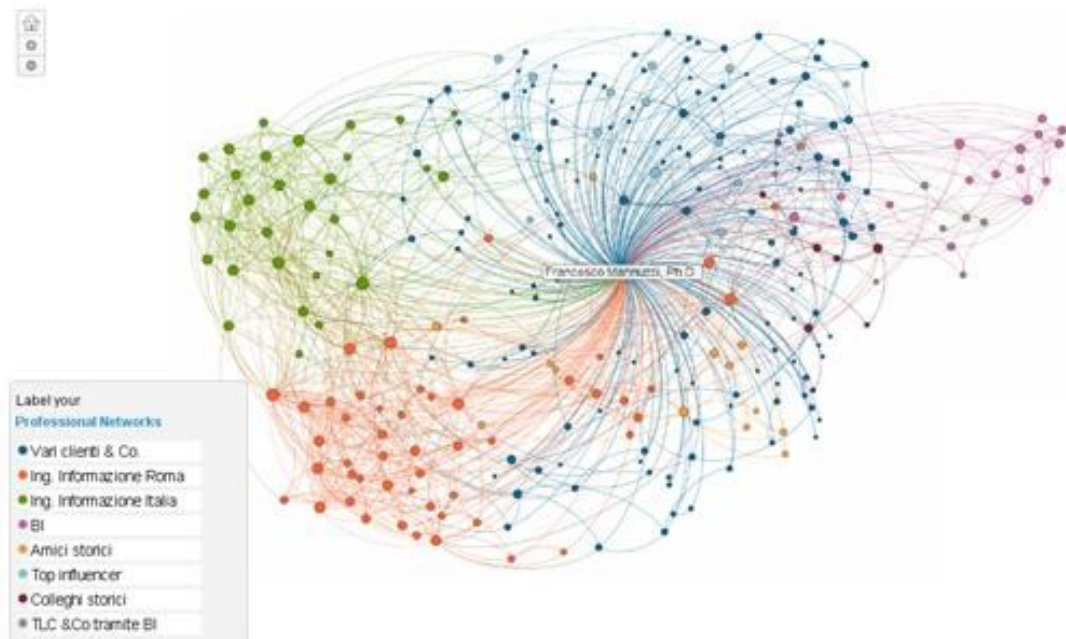
Un annuncio pubblicitario diventa effettivo quando si è in grado di predire come un utente potrebbe reagire a tale annuncio e di conseguenza indirizzare la pubblicità verso gli utenti che potrebbero reagire in modo favorevole. Nel caso della pubblicità

online si tratta di costruire un profilo degli utenti basato soprattutto sull'osservazione passiva. I Social Network Sites permettono all'utente di dichiarare i propri interessi e queste informazioni vanno a costituire una parte del profilo che le compagnie pubblicitarie cercano di costruire sui singoli individui. Un aspetto chiave dei Social Network , utile nella diffusione pubblicitaria, è la tendenza degli individui a condividere i propri interessi con gli amici e di avere la tendenza a creare amicizie con persone aventi interessi in comune. Questo aspetto risulta molto importante nel momento in cui si cerca di rispondere a domande come: gli amici hanno la tendenza a guardare annunci pubblicitari simili?.

Negli studi di marketing virale una delle conclusioni chiave è che l'adozione di un nuovo prodotto non è puramente un processo sociale. Entrano in gioco altri fattori, oltre al numero di amici, che influenzano la decisione di un utente di adottare il nuovo prodotto o meno. Richardson e Domingos hanno utilizzato un metodo simile al PageRank per stabilire quali utenti hanno maggiore influenza sulle decisioni che concernono l'adozione del nuovo prodotto. Essi scoprirono che gli individui con una maggiore influenza sugli altri utenti dovrebbero essere i primi su cui concentrare gli sforzi pubblicitari. Alcune ricerche, infatti, hanno dimostrato che il processo di adozione (del nuovo prodotto), che chiaramente coinvolge la comunicazione sociale, dipende soprattutto dalle similitudini presenti tra gli individui piuttosto che da altre componenti sociali. La ricerca del profilo di questi individui con la maggior influenza è, quindi, lo scopo delle compagnie pubblicitarie. Esse lo raggiungono raccogliendo le varie informazioni lasciate dagli utenti nel web, e principalmente estraendo informazioni dalle attività online degli utenti (ovvero i vari siti internet visitati dagli utenti) e dai Social Network.

Ad esempio di seguito è mostrato un esempio di una rete sociale delle professioni esercitate e dei collegamenti fra esse. Analizzando anche solo un semplice grafo come questo, si riesce a estrapolare numerose informazioni interessanti.

In figura è mostrato un grafo delle professioni esercitate dalle persone di una



piccola cittadina e dei collegamenti presenti fra loro.

Per poter prevedere quale annuncio pubblicitario risulta più interessante per un utente, è necessario predire quanti utenti dovrebbero effettivamente cliccare sull'annuncio. Un modello di predizione, ad esempio, si basa sulla probabilità che un utente possa cliccare su un annuncio pubblicitario appartenente a una certa classe di annunci. Gli annunci pubblicitari, infatti, vengono suddivisi in un numero finito di classi sulla base dei quali i modelli di predizione vengono sviluppati.

Questi modelli sfruttano le informazioni disponibili nei Social Network, come le preferenze dell'utente o i precedenti annunci “cliccati” dall'utente. Inoltre essi sfruttano le informazioni che si basano sugli amici dell'utente in esame. Infatti è stato dimostrato che un utente ha un'alta probabilità di rispondere a un annuncio nello stesso modo in cui hanno risposto i suoi amici. Sulla base di queste e altre informazioni contenute nei Social Network, è possibile capire quale tipologia di utente è più probabile che visualizzi un certo tipo di annuncio pubblicitario. Quest'ultima informazione, basata sul gruppo di amici, viene sfruttata anche nella predizione delle preferenze di un nuovo utente del quale non si ha nessuna

informazione sulle preferenze passate. Ovviamente le predizioni eseguite sulle informazioni raccolte dalle attività passate dell'utente sono molto più precise di quelle basate sul gruppo di amici.

In conclusione per poter eseguire una predizione migliore vengono sfruttati tutti e due i tipi di informazioni sia quelle comportamentali (basate sulle azioni passate dell'utente), sia quelle estrapolate dai Social Network (basate sul comportamento degli amici e sulle relazioni dell'utente); tutto ciò per garantire una sempre più alta probabilità di una risposta positiva agli annunci pubblicitari proposti all'utente.

2.4 Avoidance of Advertising in Social Networking Sites

La diffusione di nuovi mezzi di comunicazione ha provocato anche la diffusione di nuovi mezzi pubblicitari. I messaggi pubblicitari arrivano all'utente sotto numerose e diverse forme, come spot televisivi, finestre pop-up online, vignette sui giornali o perfino piccoli giochi su pagine internet. Come conseguenza diretta di questa massiccia diffusione di messaggi pubblicitari, gli utenti , grazie all'aiuto della tecnologia, ignorano sia i messaggi che il loro contenuto. Alcuni studi hanno dimostrato che gran parte dei consumatori sono a conoscenza del costo della pubblicità per un prodotto e di conseguenza danno più valore a prodotti poco pubblicizzati. Inoltre i consumatori sono inclini a considerare i messaggi pubblicitari più manipolativi che informativi. Considerando questi comportamenti da parte dei consumatori, è facile immaginare come questi evitino in tutti i modi possibili i messaggi pubblicitari con i quali vengono praticamente bombardati da ogni lato.

Evitare i messaggi pubblicitari vuol dire effettuare tutte quelle azioni che riducono l'esposizione del consumatore al messaggio, e possono essere azioni cognitive,

meccaniche o comportamentali. Un esempio di azione meccanica è quella di uscire dalla stanza quando viene visualizzato uno spot pubblicitario televisivo. Le nuove tecnologie facilitano il consumatore nell'evitare di essere esposto alla pubblicità, anzi gli permettono perfino di decidere come, quando e se ricevere il messaggio pubblicitario.

Schultz (nel 2006) suggerisce che questo comportamento sia una naturale reazione da parte dei consumatori a causa dell'eccessivo bombardamento pubblicitario dei mass media in via tradizionale che non. Schultz afferma che i consumatori creano delle vere e proprie barriere per evitare la pressione pubblicitaria e poter, di conseguenza, sentirsi liberi di estrapolare le informazioni ricercate su internet o su altri media, nell'esatto momento in cui ne hanno bisogno.

Studi hanno dimostrato che un fattore importante per il consumatore è la credibilità che questi ripone in uno specifico mezzo di comunicazione. Se il consumatore non ha fiducia o non crede che un mezzo di comunicazione (internet, televisione) possa essere veritiero, allora di conseguenza non presterà nemmeno attenzione agli spot pubblicitari ivi presenti. Lo stesso studio ha dimostrato che internet è il mezzo di comunicazione più credibile per la diffusione di informazioni, seguito dai quotidiani e dalla televisione. D'altra parte Internet risulta il mezzo di comunicazione meno credibile per quanto riguarda i messaggi pubblicitari, mentre i quotidiani risultano quelli più affidabili. I consumatori non si sentono al sicuro mentre navigano attraverso i vari messaggi pubblicitari presenti online perché li reputano decisamente inaffidabili. Esprimono anche il loro disagio nel comunicare online dati sensibili, come il numero della carta di credito, e tendono a effettuare acquisti online soltanto da siti che conoscono o di cui hanno fiducia.

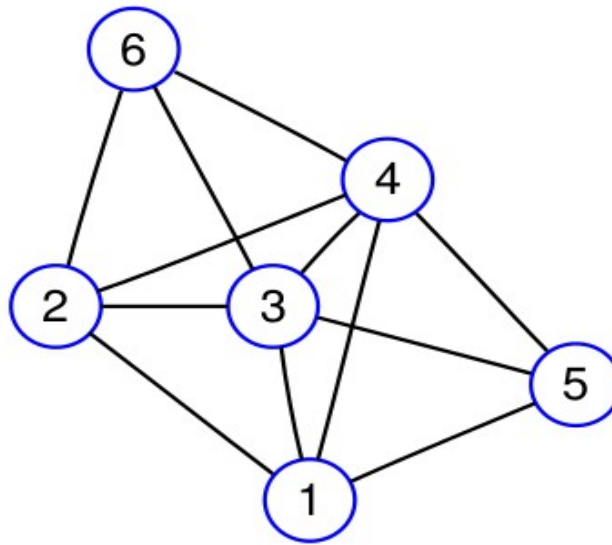
Cho e Cheon (nel 2004) hanno affermato che esistono dei precedenti che hanno sviluppato nei consumatori la sfiducia e l'abitudine di evitare i messaggi pubblicitari online. Uno di questi è denominato come : “Interruzione nel raggiungimento di un obiettivo”. Questo metodo consiste nell'inserire messaggi pubblicitari un attimo

prima che un consumatore possa raggiungere il suo obiettivo. Ad esempio quando un utente cerca di raggiungere un sito tramite un collegamento e viene interrotto in questa operazione da un messaggio pubblicitario che compare improvvisamente bloccando qualsiasi operazione. Questo tipo di pubblicità risulta fastidiosa per l'utente che non è incentivato a conoscere il contenuto pubblicitario ma, al contrario, viene incoraggiato a eliminare istantaneamente il messaggio evitandolo completamente.

Molte delle ricerche in ambito di messaggi pubblicitari online si sono concentrate su esperienze online generiche, quella al centro di questo lavoro è la rete sociale.

Fin dalla loro nascita, i Social Network hanno avuto una crescita esponenziale e il loro potenziale di raggiungere grandi masse di utenze ha suscitato l'interesse delle compagnie pubblicitarie. In generale è stato predetto che la diffusione di messaggi pubblicitari, sfruttando le reti sociali, avrebbe raggiunto i 1.2 bilioni di dollari nel 2008, e un incremento rispetto all'anno precedente del 155%.

3. The Maximum Clique Problem



In figura è mostrato un semplice esempio di grafo.

Una rete (sociale o non) viene di solito associata a un insieme di punti nello spazio, collegati tra loro da linee che rappresentano un certo tipo di relazione. Un grafo G non è altro che una rappresentazione matematica di una rete, definito da una coppia di variabili (V, E) dove V è l'insieme dei vertici ed E è l'insieme degli archi del grafo.

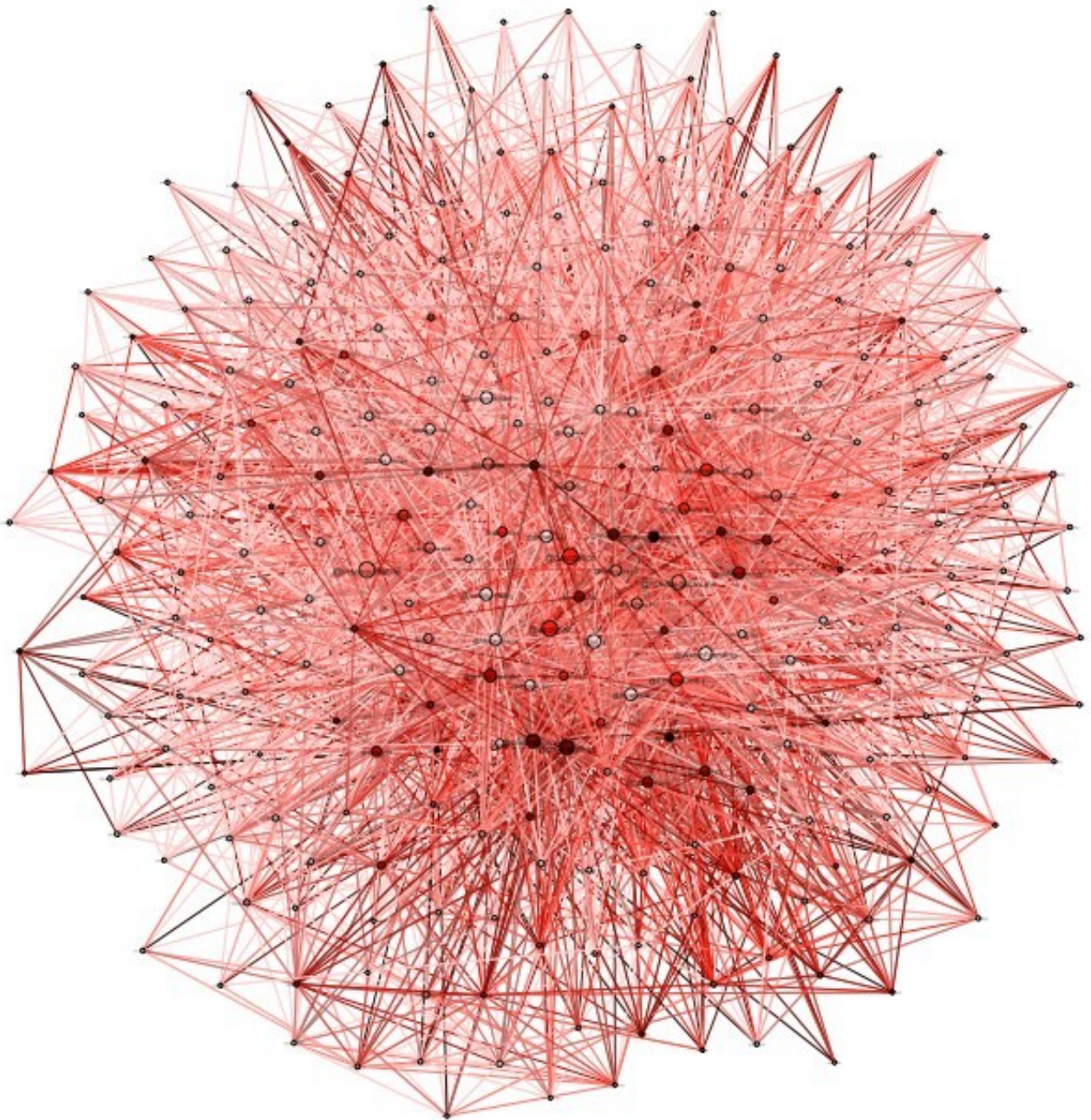
Definizione: Un grafo è una coppia ordinata $G = (V, E)$, dove V rappresenta l'insieme di nodi o vertici ed E rappresenta l'insieme degli archi tali che ogni elemento di E collega una coppia di elementi in V . Un grafo si dice:

- *orientato* quando ad ogni elemento di E è associata una coppia ordinata (u, v) con $u, v \in V$, ossia ad ogni arco è possibile associare una direzione.
- *non orientato* se agli elementi in E non è associata alcuna direzione.
- *completo* se $\forall (u, v) \in V, \exists e = (u, v) \in E$.

In questo lavoro faremo riferimento a grafi semplici non orientati.

Definizione: Un grafo $G' = (V', E')$ si dice sottografo di un grafo $G = (V, E)$ se $V' \subseteq V$ e $E' \subseteq E$.

I grafi sono degli strumenti semplici ed efficaci per modellare molte situazioni di vita reale. Generalmente i vertici rappresentano le entità mentre gli archi le relazioni tra di esse. Un esempio di un grafo viene mostrato nella figura seguente.



In figura è mostrato un esempio di un grafo con un elevato numero di nodi e archi.

Un grafo può essere di due tipi: diretto, dove ogni arco ha anche una direzione, e non diretto, dove gli archi non sono direzionati. In questo lavoro affrontiamo uno dei classici problemi sui grafi: la massima clique.

Definizione: Una clique in un grafo non orientato $G = (V, E)$ è un sottoinsieme $V' \subseteq V$ di vertici tale che ogni coppia di vertici in V' è connessa da un arco in E .

L'uso del termine clique risale a R. Duncan Luce e Perry Albert D. (A method of matrix analysis of group structure, 1949) il cui lavoro conteneva un metodo per lo studio delle relazioni sociali. Successivamente furono sviluppati algoritmi in grado di risolvere il problema della clique, il primo di questi fu presentato da Harary F. e Ross I. C. (Nel lavoro dal nome: A procedure for clique detection using the group matrix) nel 1957. Successivamente furono sviluppati numerosi algoritmi per versioni diverse del problema della massima clique. Una vera svolta in questo campo ci fu quando Cook S.A. formalizzò la nozione di riduzione polinomiale e di NP-Completezza, provando l'esistenza di un primo problema NP-Completo. Grazie a queste innovazioni di Cook, i ricercatori hanno potuto trovare giustificazioni matematiche alle difficoltà riscontrate sul problema della clique. Negli anni successivi si è giunti alla conclusione che perfino il problema di approssimare il problema della Clique è almeno NP-Completo (Lavoro documentato da: Feige U., Goldwasser S., Lovsz, L., Safra S e Szegedy M. nello scritto: Approximating clique is almost NP-complete.). Risulta quindi facilmente intuibile che anche il problema della massima clique è almeno NP-Completo (dimostrato nel lavoro: Approximating maximum independent sets by excluding subgraphs; scritto da: Boppana, R., & Halldórsson, M. nel 1992). Per questi motivi il problema della massima clique viene studiato tramite l'utilizzo di approcci euristici, come ad esempio sfruttando algoritmi greedy oppure algoritmi di local search (numerosi esempi si possono trovare nel paper: The maximum clique problem; scritto da: Bomze, I., Budinich, M., Pardalos, P., & Pelillo, M.).

Volendo definire la clique in modo diverso, essa è un sottografo completo di G . Una caratteristica importante di una clique è la sua cardinalità, ovvero il numero di vertici appartenenti alla clique stessa. Il problema della clique massima, quindi, corrisponde al problema di trovare la clique di taglia massima in un grafo. Il problema di trovare una clique di cardinalità k , invece, è un problema di decisione e consiste nel verificare se in un grafo è presente appunto una clique di taglia k . Nel lavoro di M. R. Garey & D.S. Johnson intitolato: “Computers and Intractability: A Guide to the Theory of NP-Completeness.” è stato dimostrato che, essendo il problema di decisione NP-Completo, il problema di ottimizzazione (cioè quello della massima clique) è NP-Hard.

Definizione: Una clique si dice massimale se non è sottoinsieme di nessun'altra clique.

Definizione: Una clique si dice massima se è una clique massimale con la cardinalità o il peso maggiore.

Il problema della massima clique, inoltre, risulta equivalente al problema dell'indipendent set e del vertex cover. Il primo consiste nel trovare un sottoinsieme di nodi V' del grafo G nel quale $\forall (u, v) \in V'$ non esiste $e = (u, v) \in E$. Il secondo problema, invece, consiste nel trovare il sottoinsieme V' tali che ogni arco $(i, j) \in E$ ha almeno una delle estremità in un nodo di V' .

Infine se ai nodi vengono associati dei pesi otteniamo il problema della massima clique pesata che consiste nell'individuare la clique di peso massimo nel grafo, dove il peso è dato dalla somma dei pesi dei nodi selezionati.

4. Definizioni e Modello Matematico

Di seguito verrà illustrata la formulazione del modello matematico del problema come problema di programmazione lineare intera seguito da alcuni esempi di applicazioni pratiche.

L'insieme delle variabili per la formulazione del problema della massima clique pesata è:

- variabile binaria x_i tale che per ogni $i \in V$:

$$x_i = \begin{cases} 1 & \text{se } i \text{ è selezionato} \\ 0 & \text{altrimenti} \end{cases}$$

- variabile intera w_i che rappresenta il peso del nodo i , $\forall i \in V$.

La formulazione del problema della massima clique pesata diventa:

$$\begin{aligned} \max \quad & \sum_{i=1}^n w_i x_i, \\ \text{s.t.} \quad & x_i + x_j \leq 1, \quad \forall (i, j) \in \overline{E}, \\ & x_i \in \{0, 1\}, \quad i = 1, \dots, n. \\ & w_i \geq 0 \end{aligned}$$

La funzione obiettivo della formulazione mostrata richiede di massimizzare la somma dei prodotti tra i nodi selezionati ed i rispettivi pesi.

Il primo vincolo assicura invece che due nodi x_i e x_j possono essere selezionati solo se esiste l'arco (x_i, x_j) in G . Il secondo vincolo invece assicura che la variabile x_i sia

binaria.

Infine il terzo vincolo assicura che il peso dei nodi sia un valore positivo.

Il modello studiato in questo lavoro di tesi è, invece, una variante del modello della massima clique pesata. Nella variante di seguito proposta è stata modificata la funzione obiettivo. La nuova funzione obiettivo ha come scopo principale quello di trovare la clique di peso massimo, questo vuol dire che è preferibile una soluzione con una cardinalità minore e un peso maggiore rispetto a una con una cardinalità maggiore ma un peso minore. Il modello matematico cambia come mostrato di seguito:

$$\max \quad (n + 1) \sum_{i=1}^n w_i x_i + \sum_{i=1}^n x_i$$

$$s.t. \quad x_i + x_j \leq 1, \forall (i, j) \in E$$

$$x_i \in \{0, 1\} \quad i = 1 \dots n$$

$$w_i \geq 0$$

Durante lo studio del problema sono state fatte altre conclusioni. Come detto precedentemente, l'obiettivo principale di questo lavoro è quello di raggiungere il maggior numero di utenti con gli slot pubblicitari usati. A questo scopo è stato formulato il modello matematico precedente che calcola la clique di peso massimo. Risulta interessante, anche, calcolare una clique che ha un peso di poco inferiore alla soluzione trovata, ma con una cardinalità maggiore. Una clique di questo tipo ci assicura che la distribuzione dello spot pubblicitario sia più uniforme, ovvero il numero di nodi di broadcast raggiunti è maggiore e l'informazione raggiunge quindi un numero maggiore di nodi diversi fra loro. Il modello matematico che calcola questa clique con peso più piccolo, ad esempio, di un epsilon, ma con cardinalità

maggiore, è il seguente:

$$\begin{aligned}
& \max \quad \sum_{i=1}^n x_i \\
& s.t. \quad x_i + x_j \leq 1, \forall (i, j) \in E \\
& \quad \sum_{i=1}^n w_i x_i \geq W(C^*) - \epsilon \\
& \quad \sum_{i=1}^n x_i \geq |C^*| + 1 \\
& \quad x_i \in \{0, 1\} \quad i = 1 \dots n \\
& \quad w_i \geq 0
\end{aligned}$$

La funzione obiettivo di questo modello ha come scopo quello di massimizzare la cardinalità della clique. La clique calcolata deve rispettare il vincolo che la somma dei pesi dei suoi nodi sia maggiore o uguale di $W(C^*) - \epsilon$. In questo modello $W(C^*)$ rappresenta il peso della clique di peso massimo trovata con il modello precedente ed ϵ è la tolleranza. La tolleranza rappresenta di quanto il peso della nuova clique può cambiare rispetto al peso della clique del modello precedente, nel lavoro effettuato questo parametro è stato impostato a 15% del peso della clique di massimo peso. L'ultimo vincolo indica che la clique corrente deve avere una cardinalità maggiore rispetto a quella della clique del modello precedente.

Infine è stata formulata una seconda variante di questo problema secondario in cui viene modificata la funzione obiettivo. In quest'altra variante la funzione obiettivo, oltre a massimizzare la cardinalità della clique, ha come scopo quello di massimizzare anche il peso della clique stessa.

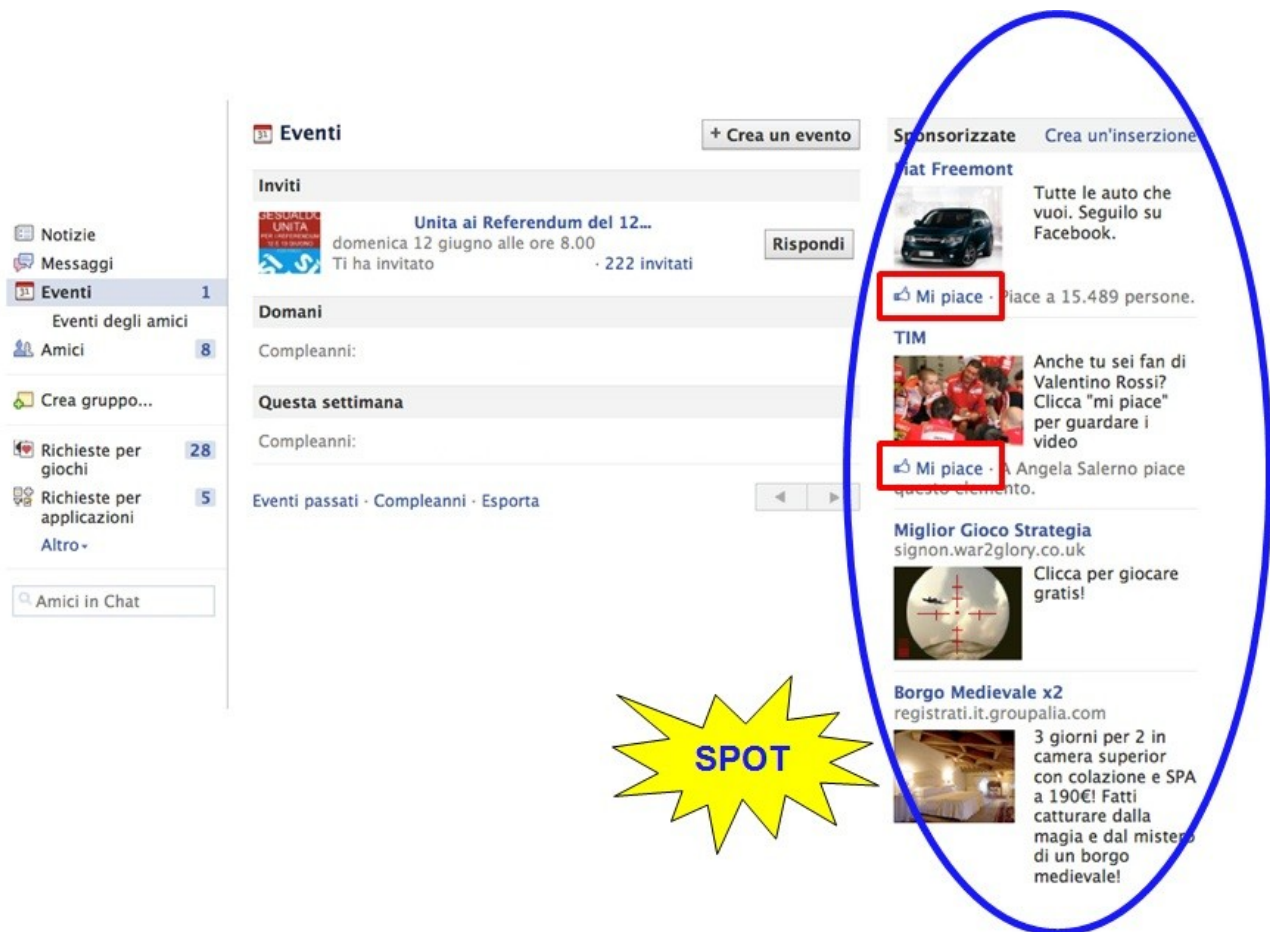
Il modello risultante è il seguente:

$$\begin{aligned}
& \max \quad W \sum_{i=1}^n x_i + \sum_{i=1}^n w_i x_i \\
& s.t. \quad x_i + x_j \leq 1, \forall (i, j) \in E \\
& \quad \sum_{i=1}^n w_i x_i \geq W(C^*) - \epsilon \\
& \quad \sum_{i=1}^n x_i \geq |C^*| + 1 \\
& \quad x_i \in \{0, 1\} \quad i = 1 \dots n \\
& \quad w_i \geq 0
\end{aligned}$$

4.1 Applicazioni Pratiche

Una delle principali applicazioni pratiche del problema studiato in questo lavoro, che suscita anche particolare interesse, è quella di massimizzare il broadcasting della pubblicità nel social network Facebook, massimizzando il numero di utenze raggiunto. L'obiettivo è quello di raggiungere il massimo numero di utenti possibile minimizzando il numero di slot pubblicitari utilizzati.

Nella figura seguente viene mostrata una tipica pagina di Facebook con in evidenza gli slot per i spot pubblicitari.



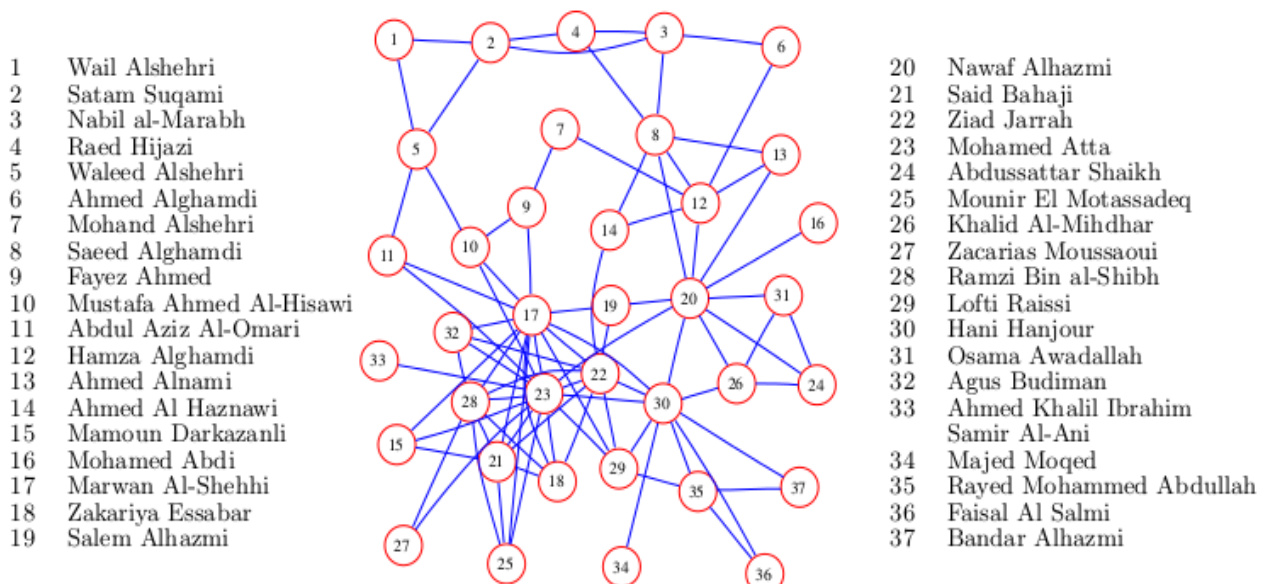
Entriamo nel dettaglio del problema. Prendiamo in considerazione l'idea di voler pubblicizzare un particolare prodotto o evento. La soluzione più ovvia e immediata sarebbe quella di inserire un annuncio pubblicitario sul profilo di ogni utente che ha tra i suoi interessi il prodotto o l'evento che si vuole pubblicizzare. Questa soluzione, anche se può sembrare efficace, in realtà non lo è affatto. Infatti viene occupato dello spazio inutilmente, spazio che potrebbe essere sfruttato per altri annunci pubblicitari. Il concetto alla base di questa affermazione è il seguente: due persone che risultano essere amiche vengono continuamente e reciprocamente informate sulle azioni effettuate dall'altro. Infatti i social network come Facebook basano il proprio servizio sulla diffusione delle informazioni, riguardanti le azioni compiute da un utente, a tutti gli amici di quello stesso utente. Questa particolare caratteristica viene sfruttata da questo lavoro nel modo seguente: si vuole sfruttare il profilo di un utente come punto di broadcast dell'annuncio pubblicitario, infatti quando un utente clicca su un “mi

piace” o “condividi”, questa informazione viene automaticamente diffusa a tutti i suoi amici. Nel nostro caso se l'utente clicca su un “mi piace” o “condividi” riguardante un annuncio pubblicitario, l'informazione di quell'annuncio viene automaticamente diffusa a tutti i suoi amici. Il problema svolto in questo lavoro è facilmente applicabile alla situazione appena descritta.

Per prima cosa si costruisce un grafo G in cui ogni nodo rappresenta un utente su Facebook e due utenti sono collegati da un arco se sono amici. Successivamente si costruisce il grafo complemento G' . Nel grafo G' ogni nodo rappresenta un gruppo di amici, ovvero è un nodo di broadcast per un particolare gruppo di amici e due nodi sono collegati da un arco se nel grafo originale G risultano sconosciuti. A ogni nodo di G' viene infine assegnato un peso che corrisponde al numero di amici rappresentati da quel particolare nodo di broadcast. Il problema quindi risulta essere trovare la massima clique pesata nel grafo complementare, ponendo particolare attenzione al peso della clique. Questo significa che la clique ricercata è preferibile che abbia un peso maggiore piuttosto che una cardinalità maggiore. Ci sono varie considerazioni da fare riguardanti questa scelta. Per prima cosa l'obiettivo principale di questo lavoro è quello di raggiungere il maggior numero di utenti, per questo motivo si preferisce ricercare la clique con il peso maggiore. Inoltre la scelta effettuata durante la costruzione del grafo complementare comporta la seguente caratteristica: due nodi di broadcast appartenenti a G' possono rappresentare gli stessi utenti. Infatti anche se due nodi del grafo G' risultano sconosciuti, questa affermazione non vale per tutti gli utenti rappresentati dai due nodi, è possibile infatti che un utente risulta essere un amico indiretto che, di conseguenza, appartiene a tutti e due gli insiemi rappresentati dai nodi di broadcast. Questa caratteristica può sembrare in un primo momento negativa perché rappresenta un tipo di ridondanza e, di conseguenza, uno spreco della diffusione della pubblicità. Dimostriamo, invece, che i vantaggi di tale situazione sono di gran lunga migliori rispetto ai svantaggi. Prendiamo, ad esempio, il caso in cui un utente di broadcast non clicca su un mi piace di un annuncio, a causa di questa scelta anche tutti gli amici di quell'utente non vengono raggiunti dalla pubblicità. Se

invece i nodi di broadcast raggruppano al loro interno alcuni stessi utenti, ovvero gli amici indiretti, allora, nella situazione descritta precedentemente, quegli stessi amici indiretti vengono ugualmente raggiunti dall'annuncio pubblicitario se un altro nodo di broadcast ha cliccato su un mi piace. Questa ridondanza, quindi, aumenta la probabilità che un gran numero di utenti venga raggiunto dallo spot pubblicitario.

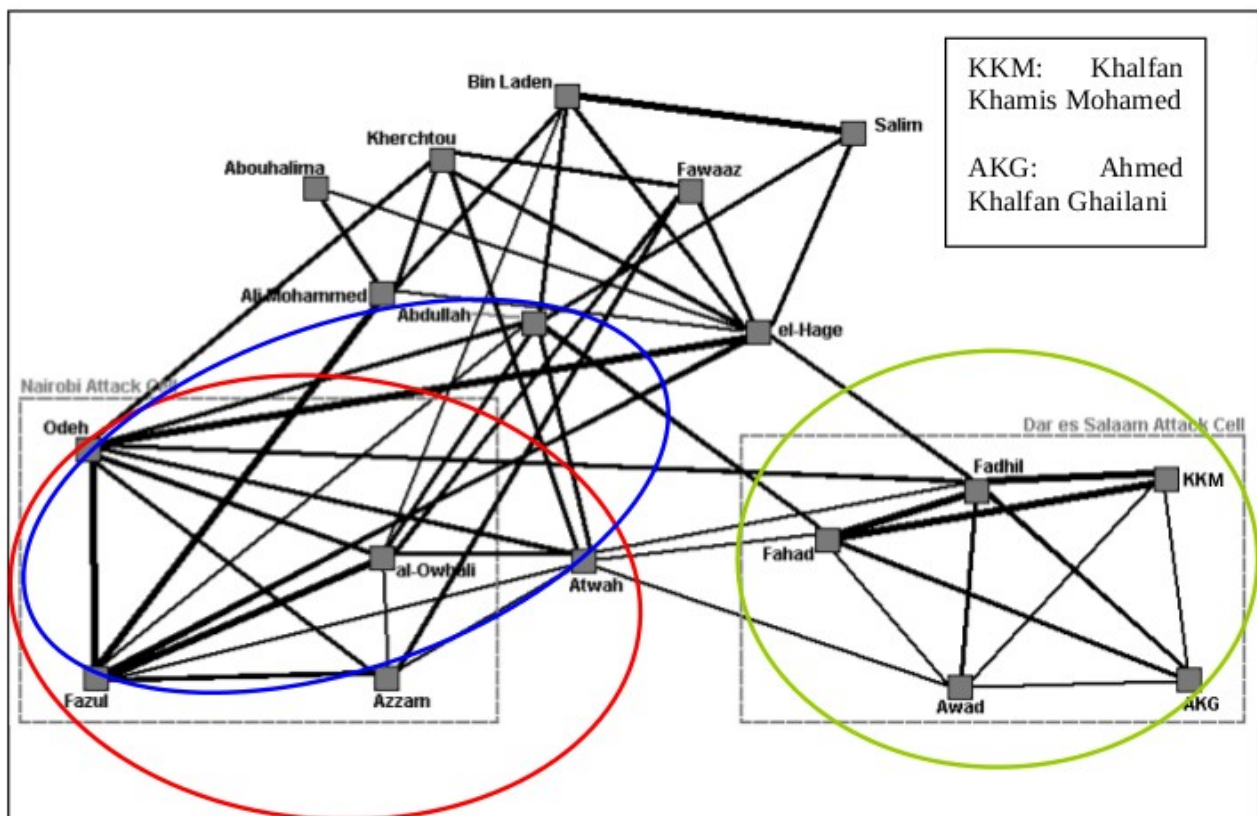
Un'altra applicazione del problema studiato si ha nel campo del Social Network Analysis. Questo è un campo nato negli ultimi dieci anni e che ha avuto notevoli sviluppi grazie all'uso della teoria dei grafi. Uno dei problemi analizzati in questo campo è la rete sociale criminale. Questo tipo di rete sociale viene rappresentata con un grafo in cui un vertice rappresenta un famoso criminale o un organizzazione criminale, e due nodi sono collegati tra loro se hanno avuto un qualche tipo di contatto che può essere un semplice scambio di informazioni o qualche tipo di affari illegali. Su questo particolare grafo si può ricercare la clique di cardinalità massima per individuare il cuore della rete sociale, ovvero l'insieme di criminali o organizzazioni che risulta essere il fulcro di transazioni illegali, eliminando il fulcro si riesce a infliggere un grosso danno alla criminalità organizzata.



In figura è rappresentata la rete terroristica dell'attentato dell'11 settembre 2001

Nel caso in cui la rete risulta essere molto grande, diventa conveniente usare un

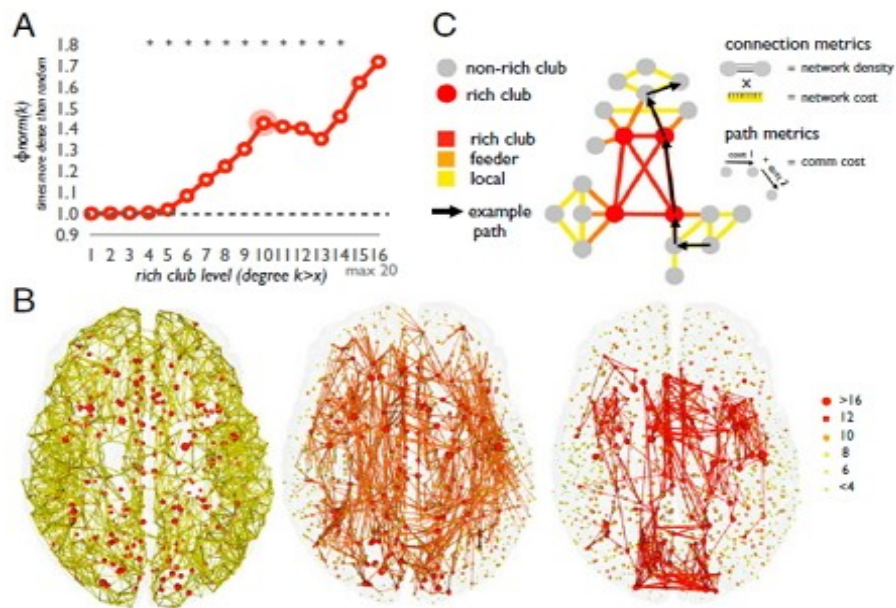
approccio diverso.



Infatti in questo caso si costruisce il grafo complementare G' come nell'esempio precedente, in cui ogni nodo è un broadcast e due nodi sono collegati tra loro se nel grafo originale erano sconosciuti.

A ogni nodo viene assegnato un peso che corrisponde a tutti i nodi con i quali il nodo di broadcast ha avuto affari illegali. Applicare il problema svolto in questo lavoro al grafo G' comporta l'individuazione della massima clique pesata. Questa clique rappresenta l'insieme di tutti i maggiori gruppi di organizzazioni o individui criminali che lavorano per conto proprio. Eliminando i nodi appartenenti a tale clique si riesce a infliggere simultaneamente un grande danno a numerosi fulcri di organizzazioni criminali. Inoltre analizzando i risultati ottenuti è possibile ricavare informazioni interessanti come, ad esempio, quali gruppi risultano sconosciuti e di conseguenza in probabile guerra fra loro.

5. Euristica e Pseudocodice



In figura è mostrato come è possibile sfruttare l'euristica DLS per trovare le zone del cervello con la maggiore attività cerebrale.

L'approccio euristico sviluppato per la risoluzione del problema è formato da una combinazione di varie fasi di espansione e di ricerca plateau in cui si cerca di trovare la massima clique possibile che rispetti i vincoli del problema, si cerca cioè di generare a ogni passo soluzioni ammissibili del problema sempre migliori.

L'approccio euristico proposto è derivato da molte altre euristiche proposte in letteratura per il problema della Massima Clique.

Nella fase principale dell'euristica lo scopo è quello di massimizzare il peso della clique inserendo vertici al suo interno ovviamente rispettando i vincoli del problema. Per questo scopo, l'algoritmo ordina un insieme di vertici considerando per ognuno: il suo grado, il grado dei vicini e peso del vicinato del nodo stesso. Successivamente inserisce i vertici all'interno della clique (inizialmente vuota) in accordo a una scelta greedy basata sul peso del vicinato controllando che il nuovo sottografo prodotto sia ancora una clique. Una volta completata l'inserzione dei vertici, si avvia la ricerca

locale per aumentare la taglia della clique corrente.

Le performance dell'algoritmo sviluppato sono state testate su un vasto insieme di istanze DIMACS, descritte successivamente. I risultati di questi test sono allegati alla fine di questo lavoro.

In questa sezione i vari passi dell'algoritmo vengono presentati in maniera più formale, mostrandone lo pseudocodice e descrivendone i particolari. L'algoritmo si basa su cinque procedure fondamentali:

- la procedura di partenza SmartClique
- la procedura Dynamic Local Search
- la procedura di espansione, chiamata Expand
- la procedura di plateau, chiamata Plateau
- la procedura di estrazione dei nodi, chiamata Extract_Heaviest_Neighbourhood.

5.1 Euristica Smart Clique

L'euristica utilizzata nella risoluzione del problema si basa principalmente su due passi :

1. Multistart
2. Dynamic Local Search

La procedura base SmartClique esegue tre principali operazioni:

1. inizializza le strutture dati necessarie;
2. implementa la strategia del multistart per garantire la stabilità dell'algoritmo;
3. chiama come subroutine le procedure che implementano i passi effettivi dell'euristica.

Lo pseudocodice della procedura è riportato di seguito.

Procedure: SmartClique

```
1: INPUT:  $G(V,E)$ , range, integer: tcs(target clique size), pd(penalty delay), maxSteps  
2: for all nodes  $\in G$   
3:   initialize Data;  
4:   Dynamic Local Search( $G, C, C1, nodes$ );  
5:end for  
6: OUTPUT: Clique Cbest in  $G$  ;
```

5.2 Multistart

Il passo di Multistart consiste semplicemente nell'eseguire il passo Dynamic Local Search tante volte quanti sono i nodi del grafo. Questa procedura risulta tanto semplice quanto importante. Da una parte incrementa notevolmente il tempo di esecuzione ma dall'altra permette una grande stabilità dell'algoritmo, caratteristica che risulta fondamentale nella ricerca della soluzione ottima desiderata. Infatti durante ogni esecuzione della procedura Dynamic Local Search viene calcolata una clique massima del grafo che viene quindi memorizzata. Durante le successive iterazioni del multistart, le clique calcolate dal DLS vengono confrontate con quella memorizzata e, nel caso la clique corrente sia migliore, la clique memorizzata viene rimpiazzata. Di conseguenza alla fine dell'esecuzione si avrà in output la miglior clique massima tra tutte quelle calcolate durante le varie esecuzioni della procedura Dynamic Local Search.

5.3 Dynamic Local Search

Il passo dell'euristica chiamato Dynamic Local Search si concentra sul calcolo di un

insieme di clique di cardinalità massima che siano ammissibili per il problema di questo lavoro.

Il passo di Dynamic Local Search è costituito a sua volta da due fasi:

- Expand Phase
- Plateau Phase

La procedura di Dynamic Local Search, insieme alla procedura di Expand e di Plateau, implementa il fulcro dell'euristica.

La procedura prende spunto dall'euristica DLS-MC presentata precedentemente; La modifica principale effettuata a quest'euristica consisteva nell'implementazione di controlli che, a ogni passo dell'esecuzione dell'algoritmo, garantissero l'ammissibilità della costruzione della soluzione. La modifica è stata necessaria in quanto l'algoritmo originale aveva come unico obiettivo il calcolo della massima clique. Più precisamente le modifiche sono state apportate nelle procedure di Expand e di Plateau chiamate a loro volta come subroutine nella procedura di Dynamic Local Search.

La Dynamic Local Search è basata, infatti, su una combinazione di fasi di espansione e di plateau search, con penalità sui vertici modificate durante la ricerca. La procedura per prima cosa azzerava un contatore che indica il numero massimo di iterazioni da eseguire prima di interrompere la ricerca di una clique migliore rispetto all'ultima trovata. Successivamente invoca la fase di Expand. Questa è una fase di miglioramento iterativo basata su un meccanismo di costruzione greedy che inizia con una clique semplice costituita da un singolo vertice e successivamente espande la clique C stessa aggiungendo vertici che sono adiacenti a tutti i vertici in C , partendo nell'espansione dal vertice con il peso del vicinato maggiore.

In questa fase si cerca di costruire iterativamente una clique di dimensione sempre più grande che continua a rispettare i vincoli del problema. Durante l'Expand per prima cosa viene costruito l'insieme dei nodi candidati a far parte della clique. Questi

nodi vengono salvati nell'insieme NIC Improving Neighbour Set of C, che è stato definito nel capitolo sulla Teoria dei Grafi. Dall'insieme NIC viene estratto il nodo avente il peso del vicinato maggiore. Questo parametro, ovvero il peso del vicinato, consiste nel sommare il peso del nodo interessato e di tutti i nodi adiacenti.

L'estrazione del nodo con il peso del vicinato maggiore rappresenta la scelta greedy dell'algoritmo sviluppato. L'obiettivo di questa politica di estrazione è quello di massimizzare il peso della clique che si va via via espandendo, con l'aggiunta di nuovi vertici. L'algoritmo continua a estrarre nodi dal NIC fino a quando questo insieme non si svuota. Quando una tale fase di espansione diventa impossibile, potrebbero ancora esistere vertici che sono connessi a tutti tranne uno dei vertici di C. Includendo tale vertice v in C e rimuovendo il singolo vertice in C non connesso a v , può essere ottenuta una nuova clique con lo stesso numero di vertici e un peso totale della clique uguale alla soluzione incombente. Questo tipo di ricerca è chiamata plateau search. A questo punto la fase di Expand termina e viene avviata la fase di Plateau. Durante la fase di Plateau si cercano soluzioni alternative ammissibili per le quali sia eventualmente possibile riavviare il processo di espansione. All'inizio della fase di Plateau viene costruito l'insieme NIL Level Neighbour Set of C, che contiene tutti i nodi che potrebbero andare a sostituire un nodo all'interno della clique mantenendo costanti le caratteristiche della soluzione, ovvero il peso e la cardinalità. Successivamente si effettua lo scambio di un nodo alla volta fino a quando l'insieme NIC risulterà non più vuoto o fino a quando l'insieme NIL risulterà non contenere altri nodi candidati. Nel primo caso viene riavviata la fase di Expand mentre nel secondo si termina l'esecuzione memorizzando la clique trovata.

Da notare che le due procedure principali, Expand e Plateau Search, sfruttano criteri diversi per la selezione del vertice da aggiungere alla clique corrente C. La procedura di Expand sfrutta la funzione `Extract_Heaviest_Neighbourhood()` per selezionare il nodo la cui somma dei pesi di tutti i nodi adiacenti è maggiore. Questo criterio di selezione greedy è stato adottato perché scegliere il nodo con il peso del vicinato maggiore implica scegliere un nodo che, con molta probabilità, porterà alla

costruzione della massima clique pesata. Nel caso della procedura Expand, la selezione viene effettuata dall'insieme NIC di tutti i vertici ammissibili che sono connessi a tutti i vertici in C; questo insieme prende il nome di Improving Neighbour set of C. Nella procedura di Plateau Search, la selezione del vertice da aggiungere alla clique corrente C, viene effettuata in base alle penalità associate a tutti i vertici candidati. Inoltre in questa procedura il vertice da aggiungere a C viene selezionato dall'insieme NLC, che prende il nome di Level Neighbour set of C. Questo insieme contiene tutti i vertici ammissibili connessi a tutti i vertici in C eccetto che per un vertice, detto v' , che viene successivamente rimosso da C.

Entrambe le procedure mantengono sempre una clique corrente C; la procedura di expand termina quando l'insieme NIC diventa vuoto, mentre la procedura di plateau termina quando o NIC non è più vuoto o NLC diventa vuoto. In più, per ridurre l'incidenza di fasi di Plateau Search non produttive, la procedura Plateau Search viene interrotta quando non c'è sovrapposizione tra la clique C' , memorizzata all'inizio del Plateau, e la clique corrente C. Alla fine della fase di plateau search, le penalità sui vertici vengono incrementate di uno. In più, ogni “pd” cicli di aggiornamento dei valori di penalità (dove pd è un parametro che significa penalty delay), tutte le penalità diverse da zero sono decrementate di uno. Quest'ultimo meccanismo viene adottato per evitare che le penalità diventino troppo grandi. Dopo questa fase di aggiornamento, la clique corrente viene perturbata in uno dei due seguenti modi. Se il valore di penalty delay “pd” è più grande di uno, la clique corrente viene ridotta all'ultimo vertice v che è stato aggiunto alla clique. Poiché tutti i vertici rimossi hanno i valori di penalità più grandi, non saranno probabilmente aggiunti nuovamente alla clique corrente nelle successive fasi di miglioramento. Questo modo di operare è equivalente a riavviare la ricerca da v . Quando il penalty delay è uno, la clique C viene perturbata aggiungendo un vertice v che viene scelto a caso in modo uniforme dal grafo in input G e rimuovendo tutti i vertici da C che non sono connessi a v . Come menzionato prima le penalità sono usate per la selezione del vertice dall'insieme NLC da aggiungere durante la procedura di Plateau Search. In

particolare la procedura che effettua la selezione si chiama `Extract_Min_Penalty()`. Quando un vertice viene selezionato tra i vicini, non sarà più disponibile per le seguenti selezioni fino a quando le penalità non vengono aggiornate e non viene eseguita la perturbazione. Queste strategie consentono all'euristica di evitare di visitare sempre la stessa clique. In più, i vertici con penalità più grande di dieci non vengono mai selezionati. Dopo ogni volta che la procedura DLS esegue una subroutine, ovvero la fase di `Expand` o quella di `Plateau Search`, effettua due controlli. Il primo controllo serve per vedere se, grazie alla subroutine, la soluzione corrente è diventata migliore di quella memorizzata fino ad ora. Il secondo controllo, invece serve per verificare se la clique corrente soddisfa i vincoli per essere scelta come la seconda massima clique. Quest'ultima clique risulta interessante perché, pur non avendo il peso massimo, essa risulta avere una cardinalità maggiore rispetto alla soluzione ottima trovata e di conseguenza una distribuzione dei pesi migliore. Applicando altre procedure sulla seconda clique massima, proposte nel capitolo sugli sviluppi futuri, potrebbe essere possibile trovare una soluzione perfino più interessante rispetto alla soluzione ottima trovata dall'algoritmo.

Di seguito è mostrato lo pseudocodice della procedura.

Procedure: Dynamic Local Search

```

1: INPUT:  $G(V,E)$ , Clique  $C$ , Clique  $C1$ , Node nodes
2: numSteps  $\leftarrow 0$ ,  $C \leftarrow \{(nodes)\}$ ;
3: while numSteps < maxSteps do
4:   if tmp nodo  $\leftarrow \text{Expand}(C, \text{true}) = \text{NULL}$  then
5:     ultimo aggiunto  $\leftarrow$  tmp nodo;
6:   end if
7:   if  $C.\text{peso} > C\text{Best}.\text{peso}$  then
8:      $C\text{Best} \leftarrow C$ ;

```

```

9:      else if C.peso == CBest.peso then
10:          if |C| > |CBest| then
11:              CBest ← C;
12:          end if
13:      end if
14:      if C.peso ≥ Cbest.peso- range && |C| > |CBest| then
15:          C2 ← C;
16:      end if
17:      C1 ← C;
18:      if tmp nodo ← P lateau(C, C1) = N U LL then
19:          ultimo aggiunto ← tmp nodo;
20:      end if
21:      while |N IC| > 0 && numSteps < maxSteps do
22:          if tmp nodo ← Expand(C, f else) = N U LL then
23:              ultimo aggiunto ← tmp nodo;
24:          end if
25:          if C.peso > CBest.peso then
26:              CBest ← C;
27:          else if C.peso == CBest.peso then
28:              if |C| > |CBest| then
29:                  CBest ← C;
30:              end if
31:          end if
32:          if C.peso ≥ Cbest.peso- range && |C| > |CBest| then
33:              C2 ← C;
34:          end if

```

```

35:         if tmp nodo  $\leftarrow$  P lateau(C, C1) = N U LL then
36:             ultimo aggiunto  $\leftarrow$  tmp nodo;
37:         end if
38:     end while
39:     UpdatePenalty(penalty delay, C);
40:     if penalty delay > 1 then
41:         C  $\leftarrow$  ultimo aggiunto;
42:     else
43:         v  $\leftarrow$  random{V };
44:         C  $\leftarrow$  C  $\cup$  {v};
45:         rimuovi tutti i vertici da C che non sono connessi a v in G;
46:     end if
47: end while
48: OUTPUT: CBest;

```

Procedura Dynamic Local Search

Procedure: Expand

```

1: INPUT: Clique C
2: Build NIC;
3: while |NIC| > 0 do
4:     temp estratto  $\leftarrow$  Extract_Heaviest_Neighbourhood(NIC);
5:     C  $\leftarrow$  C  $\cup$  {temp estratto};
6:     ultimo  $\leftarrow$  temp estratto;
7:     numSteps  $\leftarrow$  numSteps + 1;
8:     Update NIC;

```

9: **end while**

10: **OUTPUT:** nodo ultimo;

Procedura Expand

Procedure: Plateau

1: **INPUT:** Clique C, boolean built NIC

2: Build NLC;

3: **while** |NIC| = 0 && |NLC| > 0 && $C \cap C1 = \emptyset$ **do**

4: temp estratto \leftarrow selectMinPenalty(NLC);

5: da rimuovere \leftarrow selNodoDaRimuovere(C);

6: swap(da rimuovere, temp estratto);

7: numSteps \leftarrow numSteps + 1;

8: **if** C.peso > CBest.peso **then**

9: CBest \leftarrow C;

10: **else if** C.peso == CBest.peso **then**

11: **if** |C| > |CBest| **then**

12: CBest \leftarrow C;

13: **end if**

14: **end if**

15: **if** C.peso >= Cbest.peso- range && |C| > |CBest| **then**

16: C2 \leftarrow C;

17: **end if**

18: **end while**

19: ultimo \leftarrow temp estratto;

20: **OUTPUT:** nodo ultimo;

Procedura Plateau Search

6. Testing

L'euristica SmartClique è stata implementata in C++ e testata su una macchina con processore Intel Atom CPU N450 a 1.66GHz. Le istanze su cui è stato testato l'algoritmo, sono le note istanze benchmark della Second DIMACS Challenge (descritte nel lavoro dal nome: "Clique, Coloring and Satisfiability: Second DIMACS Implementation Challenge") a cui ovviamente è stata effettuata una modifica fondamentale. Le istanze originali sono costituite con specifici pesi assegnati ai nodi. Nel nostro caso, invece, le istanze devono rappresentare i grafi complementari dei grafi originali, di conseguenza i pesi dei singoli nodi sono ben specificati. Ad ogni nodo delle istanze viene assegnato un peso secondo la formula :

$$P(v) = n - \text{degree}(v)$$

Il peso " $P(v)$ " del generico nodo v è dato dal numero totale di nodi " n " del grafo meno il grado di incidenza del nodo stesso " $\text{degree}(v)$ ". Nel grafo complementare esiste un arco tra due nodi se quei nodi erano sconosciuti nel grafo originale, quindi per sapere quanti amici il generico nodo aveva nel grafo originale basta sfruttare la formula descritta precedentemente, il numero di amici corrisponde al peso del nodo nel grafo complementare.

Le famiglie di istanze su cui sono stati effettuati i test sono mostrate in tabella. La tabella è costituita da dieci colonne. La prima colonna "Id" indica degli identificativi facili per ricordare le istanze. La seconda colonna "Istanze" indica i nomi delle istanze utilizzate che fanno riferimento ai nomi originali delle istanze DIMACS. La terza e la quarta colonna indicano rispettivamente il numero di nodi e il numero di archi del grafo dell'istanza. Le colonne cinque e sei mostrano rispettivamente la cardinalità e il peso della soluzione trovata. Le colonne numero sette e otto indicano

invece la dimensione e il peso della seconda massima clique trovata. Infine l'ultima colonna indica il tempo di esecuzione dell'algoritmo su quella particolare istanza.

Ognuna delle righe delle tabelle di confronto è stata valutata come la media di tre esecuzioni distinte dell'algoritmo sulla stessa istanza .

Come si vede subito dai risultati ottenuti il tempo di esecuzione dell'algoritmo aumenta con l'aumento degli archi presenti nel grafo. Inoltre è interessante notare che in alcune particolari istanze l'algoritmo non ha calcolato la seconda massima clique, questo è semplicemente dovuto al fatto che l'algoritmo durante la costruzione del risultato ottimo non ha trovato nessuna clique che rispettasse i vincoli imposti. Aumentando il numero di iterazioni molto probabilmente queste informazioni sarebbero state calcolate.

Per la tabella dei risultati completa, composta da oltre 900 istanze, si rimanda il lettore alla fine di questo lavoro.

<i>Id</i>	<i>Istanze</i>	<i>Nodi</i>	<i>Archi</i>	<i>Size C</i>	<i>Peso C</i>	<i>Size C2</i>	<i>Peso C2</i>	<i>Time</i>
1	johnson8-2-4_c12-s11.clq	28	210	4	91	0	00	00.06 s
2	MANN_a9_c11-s15.clq	45	918	16	487	0	00	00.18 s
3	hamming6-4_c11-s27.clq	64	704	4	222	0	00	00.29 s
4	hamming6-2_c29-s43.clq	64	1824	32	1167	0	00	00.21 s
5	johnson8-4-4_c14-s51.clq	70	1855	14	625	0	00	00.25 s
6	johnson16-2-4_c23-s71.clq	120	5460	8	709	0	00	01.14 s
7	keller4_c55-s75.clq	171	9435	11	1505	9	1078	04.39 s
8	c-fat200-1_c11-s91.clq	200	1534	12	2315	0	00	01.31 s
9	c-fat200-2_c16-s107.clq	200	3235	24	4415	0	00	01.26 s
10	brock200_2_c25-s123.clq	200	9876	12	1728	11	1554	04.00 s
11	brock200_3_c60-s143.clq	200	12048	12	2142	13	1902	04.39 s
12	sanr200_0.7_c104-s167.clq	200	13868	17	2716	18	2503	05.34 s
13	san200_0.7_1_c105-s175.clq	200	13930	30	3736	20	2634	05.55 s
14	sanr200_0.9_c134-s211.clq	200	17863	41	5043	42	4452	06.22 s

15	hamming8-4_c121-s259.clq	256	20864	16	2828	0	00	06.58 s
16	hamming8-2_c185-s271.clq	256	31616	128	17087	0	00	09.59 s
17	p_hat300-1_c36-s287.clq	300	10933	8	2194	0	00	09.36 s
18	MANN_a27_c188-s323.clq	378	70551	125	24476	126	24251	43.35 s
19	san400_0.5_1_c150-s327.clq	400	39900	9	3093	9	2623	23.17 s
20	sanr400_0.5_c150-s343.clq	400	39984	12	4168	13	3792	21.24 s
21	san400_0.7_2_c70-s363.clq	400	55860	30	7974	23	6067	32.55 s
22	sanr400_0.7_c140-s395.clq	400	55869	20	6472	21	6133	30.44 s
23	brock400_3_c148-s407.clq	400	59681	25	7253	23	6597	32.55 s
24	brock400_4_c74-s427.clq	400	59765	24	7077	25	6383	33.23 s
25	johnson32-2-4_c108-s467.clq	496	107880	16	5577	0	00	74.12 s
26	c-fat500-1_c12-s479.clq	500	4459	14	6880	0	00	04.56 s
27	p_hat500-1_c94-s511.clq	500	31569	9	4145	0	00	31.02 s
28	p_hat500-3_c282-s547.clq	500	93800	48	16660	49	15150	85.40 s
29	p_hat700-1_c84-s559.clq	700	60999	11	6630	0	00	69.14 s
30	p_hat700-2_c172-s575.clq	700	121728	44	24230	44	23356	141.26 s

7. Considerazioni Finali e Sviluppi Futuri

In questo lavoro di tesi è stato introdotto un nuovo algoritmo greedy per il problema della Massima Clique Pesata. Studiando il problema sono state trovate interessanti applicazioni pratiche che spaziano dai Social Network alla Criminologia. L'esempio descritto in questo lavoro si basava sull'obiettivo di massimizzare il broadcasting della pubblicità nel social network Facebook, minimizzando al contempo il numero di slot pubblicitari utilizzati. Di questo nuovo problema è stata fornita una descrizione dettagliata ed è stato presentato un modello matematico di Programmazione Lineare. In fine è stato presentato un approccio euristico al problema. Attraverso uno schema sperimentale è stato possibile dimostrare l'efficienza effettiva della procedura confrontando i risultati del modello matematico con i risultati sperimentali. La maggior parte dello studio si è concentrata intorno all'individuazione di criteri per la rimozione di nodi e quindi anche di archi e sull'individuazione del miglior metodo per applicare la strategia greedy proposta. Inoltre lo sviluppo dell'algoritmo è stato basato sull'uso di strutture dati il più semplici e veloci possibili per migliorare le prestazioni computazionali.

Durante lo svolgimento di questo lavoro sono stati individuati interessanti spunti per successivi progetti futuri. Alcuni di questi sono i seguenti:

- Individuare ulteriori criteri di rimozione e rendere ulteriormente efficienti quelli attuali
- Studiare e analizzare la fase di Plateau in modo da poter effettuare lo scambio di due nodi anche quando quest'operazione comporta una soluzione peggiore dell'attuale ma permette di espandere ulteriormente la clique, dando la possibilità di trovare una soluzione altrimenti scartata a priori.
- Studiare e analizzare nuovi metodi per eliminare dal grafo nodi che non

potrebbero mai far parte della soluzione finale, come ad esempio il calcolo del peso del vicinato di un nodo. Infatti se il peso del vicinato di un nodo è inferiore al peso della clique corrente, allora sicuramente non potrà far parte della soluzione finale, quindi è possibile eliminarlo per semplificare e velocizzare la ricerca della soluzione ottima.

- Studiare e implementare procedure che sfruttano alcuni dei risultati ottenuti dall'algoritmo sviluppato in questo lavoro. Ad esempio è possibile implementare una procedura che, partendo dalla seconda clique massima trovata, applichi procedure come la Plateau Search con l'unico scopo di aumentare la cardinalità della clique, ignorando gli altri vincoli del problema. Questa applicazione risulta molto interessante in quanto il peso della clique calcolata risulta inferiore a quella ottima ma la sua distribuzione risulta migliore.

8. Bibliografia

1. A graph theoretical approach for predicting common RNA secondary structure motifs including pseudoknots in unaligned sequences. Ji, Y., Xu, & G. D. Stormo , *Bioinformatics*, 20(10), 1591-1602, (2004).
2. A method of matrix analysis of group structure, Luce R. Duncan e Perry Albert D., *Psychometrika*, 14 (2), 95-116,(1949).
3. A new Table of Constant Weight Codes A. E. Brouwer, J. B. Shearer, N. J. A. Sloane and W. D. Smith, *J. IEEE Trans. Information Theory*, 36, 1334-1380, (1990).
4. A new trust region technique for the maximum clique problem. S. Busygin , Internal report, <http://www.busygin.dp.ua>. (2002).
5. A procedure for clique detection using the group matrix, Harary F. e Ross I. C. *Sociometry (American Sociological Association)* 20 (3), 205-215,(1957).
6. A study of algorithm capabilities for solving the maximum clique problem. C. Solnon & S. Fenet, *Journal of Heuristics*, to appear.
7. Additive versus multiplicative clause weighting for SAT. J. Thornton, D. N. Pham, S. Bain & V. Ferreira , In *Proceedings of the 19th National Conference on Artificial Intelligence, (AAAI-04)*, 191-196, AAAI Press / The MIT Press, Menlo Park, CA, USA, (2004).
8. An effective local search for the maximum clique problem. K. Katayama, A. Hamamoto, H. Narihisa , *Information Processing Letters*, 95 Issue 5, (2005) .
9. Approximating clique is almost NP-complete. Feige U., Goldwasser S., Lovsz, L., Safra S e Szegedy M. *Proc. 32nd IEEE Symp. on Foundations of Computer Science*, 2-12, (1991).

10. Approximating maximum independent sets by o
11. Approximation Algorithms. Vijay V. Vazirani Springer, (2003).
12. Clique is hard to approximate within $n^{1-\epsilon}$. Hastad, J. Acta Mathematica, 182, 105- 142, (1999).
13. Clique, Coloring and Satisfiability: Second DIMACS Implementation Challenge. D. Johnson & M. Trick (Eds.). , DIMACS Series. American Mathematical Society, Vol. 26, (1996).
14. Combinatorial approaches to finding subtle signals in DNA sequences. P.A. Pevzner & S.-H. Sze , In Proceedings of the Eighth International Conference on Intelligent Systems for Molecular Biology, AAAI Press, 269-278, (2000).
15. Combining swaps and node weights in an adaptive greedy approach for the maximum clique problem. A. Grosso, M. Locatelli, F.D. Croce, , Journal of Heuristics, 10, 135-152, (2004).
16. Complete subgraphs make an early appearance in the mathematical literature in the graph-theoretic reformulation, Erdos e Szekeres, Ramsey theory, (1935).
17. Computers and Intractability: A Guide to the Theory of NP-Completeness. M. R. Garey & D.S. Johnson , Freeman, San Francisco, CA, USA, (1979).
18. Dynamic Local Search for the Maximum Clique Problem. W. Pullan & H.H. Hoos , Journal of Artificial Intelligence Research, 25, 159-185, (2006).
19. excluding subgraphs. Boppana, R., & Halldórsson, M. Bit, 32, 180-196, (1992).
20. Extensions of the minimum labelling spanning tree problem. R. Cerulli, A. Fink, M. Gentili, & S. Voss. Journal of Telecommunication and Information Technology, 4, 39-45, (2006).
21. Finding a maximum clique in an arbitrary graph. E. Balas & C. Yu , SIAM Journal of Computing, 15(4), 1054-1068, (1986).

22. Finding a maximum independent set. Tarjan R. E. e Trojanowski A. E. SIAM Journal on Computing 6 (3), 537-546, (1997).
23. Math Enters Age of Electronic Mail. Kolata Gina In a Frenzy, New York Times, (1990).
24. Metaheuristics comparison for the minimum labelling spanning tree problem. R. Cerulli, A. Fink, M. Gentili, & S. Voss. In B. Golden, S. Raghavan, and E. Wasil, editors, The Next Wave in Computing, Optimization, and Decision Technologies, 93-106. Springer, (2005).
25. On the red-blue set cover problem. R.D. Carr, S. Doddi, G. Konjedov, & M. Marathe. In 11th ACM-SIAM Symposium on Discrete Algorithms, 345-353, (2000).
26. Reactive Local Search for the Maximum Clique Problem: a new implementation. R. Battiti & F. Mascia , Technical Report, (2007).
27. Reactive Local Search for the Maximum Clique Problem. R. Battiti e M. Protasi , Algorithmica, 29, 610-637, (2001).
28. Reducibility among combinatorial problems, R. Karp, Plenum Press, (1972).
29. Resolvent clause weighting local search. W. Pullan & L. Zhao , In Tawfik, A. Y. & Goodwin, S. D. (Eds.), Advances in Artificial Intelligence, 17th Conference of the Canadian Society for Computational Studies of Intelligence, Vol. 3060 of Lecture Notes in Computer Science, 233-247, Springer Verlag, Berlin, Germany, (2004).
30. Scaling and probabilistic smoothing: Dynamic local search for unweighted MAX-SAT. D. Tompkins & H. Hoos , In Xiang, Y., & Chaib-draa, B. (Eds.), Advances in Artificial Intelligence, 16th Conference of the Canadian Society for Computational Studies of Intelligence, Vol. 2671 of Lecture Notes in Computer Science, 145-159, Springer Verlag, Berlin, Germany, (2003).
31. Scaling and probabilistic smoothing: efficient dynamic local search for SAT. F.

- Hutter, D. A. D. Tompkins & H. H. Hoos , In Hentenryck, P.V. (Ed.), Principles and Practice of Constraint Programming -CP, Vol. 2470 of Lecture Notes in Computer Science, Springer Verlag, Berlin, Germany, 233-248, (2002).
32. Searching for Maximum Cliques with Ant Colony Optimization. S. Fenet & C. Solnon , Lecture Notes in Computer Science, Volume 2611/2003, 291-302, (2003) .
33. Social network analysis: methods and applications. S. Wasserman, K. Faust. , Cambridge University Press, (1999).
34. Solving the maximum clique problem by k-opt local search. K. Katayama, A. Hamamoto, H. Narihisa , In proceedings of the 2004 ACM Symposium on Applied computing, 1021-1025, (2004).
35. Solving the maximum clique problem by k-opt local search., K. Katayama, A. Hamamoto, H. Narihisa. , In Proceedings of of the 2004 ACM Symposium on Applied computing, 1021-1025, (2004).
36. Spanning trees with many or few colors in edge-colored graphs. H. Broersma & X. Li. *Discussiones Mathematicae Graph Theory*, 17(2), 259-269, (1997).
37. The complexity of theorem-proving procedures. Cook S. A. *Proc. 3rd ACM Symposium on Theory of Computing*, 151-158, (1971).
38. The Labeled Maximum Matching Problem. F. Carrabs, R. Cerulli, M. Gentili. *Computers & Operations Research*, 36(6), 1859-1871, (2009).
39. The maximum clique problem. Bomze, I., Budinich, M., Pardalos, P., & Pelillo, M. In D.Z. Du, P.P.(Ed.). *Handbook of Combinatorial Optimization*, Vol. A, 1- 4, (1999).
40. The Maximum Clique Problem. P.M. Pardalos e J. Xue , *Journal of Global Optimization*, 4(3) 301-328, (1992).

41. The minimum labeling spanning trees. R.S. Chang & S.J. Leu. Information Processing Letters, 63(5), 277-282, (1997).
42. The Reactive Tabu Search. R. Battiti e G. Tecchiolli. ORSA Journal on Computing, 6(2), 126-140,(1994).
43. Topic on Perfect Graph. C. Berge, V.Chvatal , Ann Discr. Math., 21, (1984).
44. Universal sequential search problems. Levin L. A. Problems of Information Transmission, 9 (3), 265-266, (1973).

9. Tabella Risultati Testing

Istanze	Nodi	Archi	Size C	Peso C	Size C2	Peso C2	Time
johnson8-2-4_c12-s11.clq	28	210	4	91	0	00	00.06 s
johnson8-2-4_c12-s3.clq	28	210	4	91	0	00	00.06 s
johnson8-2-4_c12-s7.clq	28	210	4	91	0	00	00.06 s
johnson8-2-4_c4-s3.clq	28	210	4	91	0	00	00.06 s
johnson8-2-4_c4-s3.clq	28	210	4	91	0	00	00.06 s
johnson8-2-4_c4-s3.clq	28	210	4	91	0	00	00.06 s
johnson8-2-4_c8-s3.clq	28	210	4	91	0	00	00.06 s
johnson8-2-4_c8-s3.clq	28	210	4	91	0	00	00.06 s
johnson8-2-4_c8-s3.clq	28	210	4	91	0	00	00.06 s
MANN_a9_c11-s15.clq	45	918	16	487	0	00	00.18 s
MANN_a9_c11-s19.clq	45	918	16	487	0	00	00.16 s
MANN_a9_c11-s23.clq	45	918	16	487	0	00	00.17 s
MANN_a9_c21-s15.clq	45	918	16	490	0	00	00.17 s
MANN_a9_c21-s19.clq	45	918	16	487	0	00	00.16 s
MANN_a9_c21-s23.clq	45	918	16	487	0	00	00.16 s
MANN_a9_c32-s15.clq	45	918	16	490	0	00	00.17 s
MANN_a9_c32-s19.clq	45	918	16	487	0	00	00.16 s
MANN_a9_c32-s23.clq	45	918	16	490	0	00	00.17 s
hamming6-4_c11-s27.clq	64	704	4	222	0	00	00.29 s
hamming6-4_c11-s31.clq	64	704	4	222	0	00	00.28 s
hamming6-4_c11-s35.clq	64	704	4	222	0	00	00.29 s
hamming6-4_c17-s27.clq	64	704	4	222	0	00	00.29 s
hamming6-4_c17-s31.clq	64	704	4	222	0	00	00.28 s
hamming6-4_c17-s35.clq	64	704	4	222	0	00	00.29 s
hamming6-4_c6-s27.clq	64	704	4	222	0	00	00.28 s
hamming6-4_c6-s31.clq	64	704	4	222	0	00	00.28 s
hamming6-4_c6-s35.clq	64	704	4	222	0	00	00.29 s
hamming6-2_c15-s39.clq	64	1824	32	1167	0	00	00.22 s
hamming6-2_c15-s43.clq	64	1824	32	1167	0	00	00.22 s

hamming6-2_c15-s47.clq	64	1824	32	1167	0	00	00.21 s
hamming6-2_c29-s39.clq	64	1824	32	1167	0	00	00.22 s
hamming6-2_c29-s43.clq	64	1824	32	1167	0	00	00.21 s
hamming6-2_c29-s47.clq	64	1824	32	1167	0	00	00.22 s
hamming6-2_c44-s39.clq	64	1824	32	1167	0	00	00.22 s
hamming6-2_c44-s43.clq	64	1824	32	1167	0	00	00.22 s
hamming6-2_c44-s47.clq	64	1824	32	1167	0	00	00.22 s
johnson8-4-4_c14-s51.clq	70	1855	14	625	0	00	00.25 s
johnson8-4-4_c14-s55.clq	70	1855	14	625	0	00	00.26 s
johnson8-4-4_c14-s59.clq	70	1855	14	625	0	00	00.26 s
johnson8-4-4_c27-s51.clq	70	1855	14	625	0	00	00.26 s
johnson8-4-4_c27-s55.clq	70	1855	14	625	0	00	00.27 s
johnson8-4-4_c27-s59.clq	70	1855	14	625	0	00	00.26 s
johnson8-4-4_c40-s51.clq	70	1855	14	625	0	00	00.26 s
johnson8-4-4_c40-s55.clq	70	1855	14	625	0	00	00.26 s
johnson8-4-4_c40-s59.clq	70	1855	14	625	0	00	00.26 s
johnson16-2-4_c23-s63.clq	120	5460	8	709	0	00	01.15 s
johnson16-2-4_c23-s67.clq	120	5460	8	709	0	00	01.16 s
johnson16-2-4_c23-s71.clq	120	5460	8	709	0	00	01.14 s
johnson16-2-4_c46-s63.clq	120	5460	8	709	0	00	01.14 s
johnson16-2-4_c46-s67.clq	120	5460	8	709	0	00	01.13 s
johnson16-2-4_c46-s71.clq	120	5460	8	709	0	00	01.16 s
johnson16-2-4_c69-s63.clq	120	5460	8	709	0	00	01.17 s
johnson16-2-4_c69-s67.clq	120	5460	8	709	0	00	01.15 s
johnson16-2-4_c69-s71.clq	120	5460	8	709	0	00	01.14 s
keller4_c28-s75.clq	171	9435	11	1505	9	1078	04.40 s
keller4_c28-s79.clq	171	9435	11	1505	9	1078	04.39 s
keller4_c28-s83.clq	171	9435	11	1505	9	1078	04.39 s
keller4_c55-s75.clq	171	9435	11	1505	9	1078	04.39 s
keller4_c55-s79.clq	171	9435	11	1505	9	1078	04.40 s
keller4_c55-s83.clq	171	9435	11	1505	9	1078	04.39 s
keller4_c83-s75.clq	171	9435	11	1505	9	1078	04.42 s
keller4_c83-s79.clq	171	9435	11	1505	9	1078	04.40 s
keller4_c83-s83.clq	171	9435	11	1505	9	1078	04.40 s
c-fat200-1_c11-s87.clq	200	1534	12	2315	0	00	01.30 s
c-fat200-1_c11-s91.clq	200	1534	12	2315	0	00	01.31 s

c-fat200-1_c11-s95.clq	200	1534	12	2315	0	00	01.30 s
c-fat200-1_c4-s87.clq	200	1534	12	2315	0	00	01.29 s
c-fat200-1_c4-s91.clq	200	1534	12	2315	0	00	01.30 s
c-fat200-1_c4-s95.clq	200	1534	12	2315	0	00	01.30 s
c-fat200-1_c7-s87.clq	200	1534	12	2315	0	00	01.31 s
c-fat200-1_c7-s91.clq	200	1534	12	2315	0	00	01.30 s
c-fat200-1_c7-s95.clq	200	1534	12	2315	0	00	01.31 s
c-fat200-2_c16-s103.clq	200	3235	24	4415	0	00	01.27 s
c-fat200-2_c16-s107.clq	200	3235	24	4415	0	00	01.26 s
c-fat200-2_c16-s99.clq	200	3235	24	4415	0	00	01.27 s
c-fat200-2_c24-s103.clq	200	3235	24	4415	0	00	01.26 s
c-fat200-2_c24-s107.clq	200	3235	24	4415	0	00	01.28 s
c-fat200-2_c24-s99.clq	200	3235	24	4415	0	00	01.29 s
c-fat200-2_c8-s103.clq	200	3235	24	4415	0	00	01.26 s
c-fat200-2_c8-s107.clq	200	3235	24	4415	0	00	01.28 s
c-fat200-2_c8-s99.clq	200	3235	24	4415	0	00	01.26 s
c-fat200-5_c21-s111.clq	200	8473	58	9192	0	00	03.20 s
c-fat200-5_c21-s115.clq	200	8473	58	9192	0	00	03.18 s
c-fat200-5_c21-s119.clq	200	8473	58	9192	0	00	03.19 s
c-fat200-5_c42-s111.clq	200	8473	58	9192	0	00	03.22 s
c-fat200-5_c42-s115.clq	200	8473	58	9192	0	00	03.21 s
c-fat200-5_c42-s119.clq	200	8473	58	9192	0	00	03.19 s
c-fat200-5_c63-s111.clq	200	8473	58	9192	0	00	03.18 s
c-fat200-5_c63-s115.clq	200	8473	58	9192	0	00	03.18 s
c-fat200-5_c63-s119.clq	200	8473	58	9192	0	00	03.21 s
brock200_2_c25-s123.clq	200	9876	12	1728	11	1554	04.00 s
brock200_2_c25-s127.clq	200	9876	12	1728	10	1577	04.00 s
brock200_2_c25-s131.clq	200	9876	12	1728	11	1525	04.00 s
brock200_2_c49-s123.clq	200	9876	12	1728	11	1525	04.37 s
brock200_2_c49-s127.clq	200	9876	12	1728	10	1553	04.39 s
brock200_2_c49-s131.clq	200	9876	12	1728	11	1525	04.39 s
brock200_2_c74-s123.clq	200	9876	12	1728	11	1554	04.37 s
brock200_2_c74-s127.clq	200	9876	12	1728	11	1525	04.39 s
brock200_2_c74-s131.clq	200	9876	12	1728	11	1554	04.37 s
brock200_3_c30-s135.clq	200	12048	12	2142	13	1902	04.39 s
brock200_3_c30-s139.clq	200	12048	12	2142	13	1853	04.40 s

brock200_3_c30-s143.clq	200	12048	12	2142	14	2116	04.59 s
brock200_3_c60-s135.clq	200	12048	12	2142	13	1798	04.59 s
brock200_3_c60-s139.clq	200	12048	12	2142	13	1995	04.37 s
brock200_3_c60-s143.clq	200	12048	12	2142	13	1902	04.39 s
brock200_3_c90-s135.clq	200	12048	12	2142	13	1716	04.43 s
brock200_3_c90-s139.clq	200	12048	12	2142	13	1815	05.00 s
brock200_3_c90-s143.clq	200	12048	12	2142	13	1718	05.00 s
brock200_4_c33-s147.clq	200	13089	15	2532	16	2280	05.30 s
brock200_4_c33-s151.clq	200	13089	15	2532	16	2460	05.29 s
brock200_4_c33-s155.clq	200	13089	15	2532	16	2477	05.28 s
brock200_4_c65-s147.clq	200	13089	15	2532	16	2477	05.29 s
brock200_4_c65-s151.clq	200	13089	15	2532	16	2479	05.28 s
brock200_4_c65-s155.clq	200	13089	15	2532	16	2460	05.27 s
brock200_4_c98-s147.clq	200	13089	15	2532	16	2280	05.29 s
brock200_4_c98-s151.clq	200	13089	15	2532	16	2468	05.26 s
brock200_4_c98-s155.clq	200	13089	15	2532	16	2477	05.27 s
sanr200_0.7_c104-s159.clq	200	13868	17	2716	18	2526	05.37 s
sanr200_0.7_c104-s163.clq	200	13868	17	2716	18	2486	05.35 s
sanr200_0.7_c104-s167.clq	200	13868	17	2716	18	2503	05.34 s
sanr200_0.7_c35-s159.clq	200	13868	17	2716	18	2537	05.34 s
sanr200_0.7_c35-s163.clq	200	13868	17	2716	18	2397	05.34 s
sanr200_0.7_c35-s167.clq	200	13868	17	2716	18	2477	05.34 s
sanr200_0.7_c69-s159.clq	200	13868	17	2716	18	2537	05.33 s
sanr200_0.7_c69-s163.clq	200	13868	17	2716	18	2477	05.38 s
sanr200_0.7_c69-s167.clq	200	13868	17	2716	18	2503	05.35 s
san200_0.7_1_c105-s171.clq	200	13930	30	3736	21	2623	05.50 s
san200_0.7_1_c105-s175.clq	200	13930	30	3736	20	2634	05.55 s
san200_0.7_1_c105-s179.clq	200	13930	30	3736	19	2417	05.42 s
san200_0.7_1_c35-s171.clq	200	13930	30	3736	20	2577	05.39 s
san200_0.7_1_c35-s175.clq	200	13930	30	3736	21	2500	05.46 s
san200_0.7_1_c35-s179.clq	200	13930	30	3736	20	2561	05.55 s
san200_0.7_1_c70-s171.clq	200	13930	30	3736	20	2635	06.00 s
san200_0.7_1_c70-s175.clq	200	13930	30	3736	20	2508	05.43 s
san200_0.7_1_c70-s179.clq	200	13930	30	3736	18	2539	05.48 s
san200_0.7_2_c105-s183.clq	200	13930	14	2379	14	2296	05.23 s
san200_0.7_2_c105-s187.clq	200	13930	13	2373	14	1960	05.24 s

san200_0.7_2_c105-s191.clq	200	13930	14	2379	15	1977	05.22 s
san200_0.7_2_c35-s183.clq	200	13930	14	2413	15	2312	05.43 s
san200_0.7_2_c35-s187.clq	200	13930	13	2400	14	2162	05.48 s
san200_0.7_2_c35-s191.clq	200	13930	14	2372	15	1985	05.50 s
san200_0.7_2_c70-s183.clq	200	13930	14	2372	15	2187	05.55 s
san200_0.7_2_c70-s187.clq	200	13930	18	2484	16	2231	05.42 s
san200_0.7_2_c70-s191.clq	200	13930	13	2373	14	2296	05.22 s
brock200_1_c111-s195.clq	200	14834	19	3033	20	2462	05.26 s
brock200_1_c111-s199.clq	200	14834	19	3033	20	2684	05.24 s
brock200_1_c111-s203.clq	200	14834	19	3033	20	2574	05.25 s
brock200_1_c37-s195.clq	200	14834	19	3033	20	2439	05.22 s
brock200_1_c37-s199.clq	200	14834	19	3033	20	2579	05.23 s
brock200_1_c37-s203.clq	200	14834	19	3033	20	2479	05.24 s
brock200_1_c74-s195.clq	200	14834	19	3033	20	2674	05.28 s
brock200_1_c74-s199.clq	200	14834	19	3033	20	2585	05.23 s
brock200_1_c74-s203.clq	200	14834	19	3033	20	2684	05.27 s
sanr200_0.9_c134-s207.clq	200	17863	41	5043	42	4467	06.22 s
sanr200_0.9_c134-s211.clq	200	17863	41	5043	42	4452	06.22 s
sanr200_0.9_c134-s215.clq	200	17863	41	5043	42	4467	06.20 s
sanr200_0.9_c45-s207.clq	200	17863	41	5043	42	4452	06.16 s
sanr200_0.9_c45-s211.clq	200	17863	41	5043	42	4452	06.22 s
sanr200_0.9_c45-s215.clq	200	17863	41	5043	42	4452	06.17 s
sanr200_0.9_c89-s207.clq	200	17863	41	5043	42	4452	06.18 s
sanr200_0.9_c89-s211.clq	200	17863	41	5043	42	4452	06.20 s
sanr200_0.9_c89-s215.clq	200	17863	41	5043	42	4467	06.19 s
san200_0.9_1_c135-s219.clq	200	17910	70	7879	54	5797	06.46 s
san200_0.9_1_c135-s223.clq	200	17910	70	7879	54	5797	06.46 s
san200_0.9_1_c135-s227.clq	200	17910	70	7879	54	5797	06.43 s
san200_0.9_1_c45-s219.clq	200	17910	70	7879	54	5797	06.43 s
san200_0.9_1_c45-s223.clq	200	17910	70	7879	54	5797	06.42 s
san200_0.9_1_c45-s227.clq	200	17910	70	7879	53	5876	06.44 s
san200_0.9_1_c90-s219.clq	200	17910	70	7879	54	5797	06.46 s
san200_0.9_1_c90-s223.clq	200	17910	70	7879	54	5797	06.43 s
san200_0.9_1_c90-s227.clq	200	17910	70	7879	54	5797	06.46 s
san200_0.9_2_c135-s231.clq	200	17910	60	6699	41	4371	06.59 s
san200_0.9_2_c135-s235.clq	200	17910	60	6699	44	5180	06.58 s

san200_0.9_2_c135-s239.clq	200	17910	60	6699	44	5180	06.57 s
san200_0.9_2_c45-s231.clq	200	17910	60	6699	44	5180	06.57 s
san200_0.9_2_c45-s235.clq	200	17910	60	6699	44	5180	06.57 s
san200_0.9_2_c45-s239.clq	200	17910	60	6699	44	5180	06.58 s
san200_0.9_2_c90-s231.clq	200	17910	60	6699	46	5180	07.00 s
san200_0.9_2_c90-s235.clq	200	17910	60	6699	44	5180	06.58 s
san200_0.9_2_c90-s239.clq	200	17910	60	6699	44	5180	06.57 s
san200_0.9_3_c135-s243.clq	200	17910	36	4934	37	4013	06.50 s
san200_0.9_3_c135-s247.clq	200	17910	32	4863	37	4013	07.00 s
san200_0.9_3_c135-s251.clq	200	17910	35	4850	36	4023	06.52 s
san200_0.9_3_c45-s243.clq	200	17910	33	4907	35	3954	06.59 s
san200_0.9_3_c45-s247.clq	200	17910	35	4844	44	4797	06.49 s
san200_0.9_3_c45-s251.clq	200	17910	36	4934	44	4797	06.47 s
san200_0.9_3_c90-s243.clq	200	17910	32	4893	35	3917	07.13 s
san200_0.9_3_c90-s247.clq	200	17910	36	4868	37	4013	06.55 s
san200_0.9_3_c90-s251.clq	200	17910	33	4875	35	3909	06.58 s
hamming8-4_c121-s255.clq	256	20864	16	2828	0	00	06.52 s
hamming8-4_c121-s259.clq	256	20864	16	2828	0	00	06.58 s
hamming8-4_c121-s263.clq	256	20864	16	2828	0	00	07.00 s
hamming8-4_c41-s255.clq	256	20864	16	2828	0	00	06.52 s
hamming8-4_c41-s259.clq	256	20864	16	2828	0	00	06.57 s
hamming8-4_c41-s263.clq	256	20864	16	2828	0	00	06.52 s
hamming8-4_c81-s255.clq	256	20864	16	2828	0	00	07.00 s
hamming8-4_c81-s259.clq	256	20864	16	2828	0	00	06.56 s
hamming8-4_c81-s263.clq	256	20864	16	2828	0	00	06.58 s
hamming8-2_c123-s267.clq	256	31616	128	17087	0	00	09.59 s
hamming8-2_c123-s271.clq	256	31616	128	17087	0	00	09.58 s
hamming8-2_c123-s275.clq	256	31616	128	17087	0	00	09.54 s
hamming8-2_c185-s267.clq	256	31616	128	17087	0	00	09.59 s
hamming8-2_c185-s271.clq	256	31616	128	17087	0	00	09.59 s
hamming8-2_c185-s275.clq	256	31616	128	17087	0	00	10.00 s
hamming8-2_c62-s267.clq	256	31616	128	17087	0	00	06.58 s
hamming8-2_c62-s271.clq	256	31616	128	17087	0	00	09.53 s
hamming8-2_c62-s275.clq	256	31616	128	17087	0	00	09.58 s
p_hat300-1_c18-s279.clq	300	10933	8	2194	0	00	09.36 s
p_hat300-1_c18-s283.clq	300	10933	8	2194	0	00	09.32 s

p_hat300-1_c18-s287.clq	300	10933	8	2194	0	00	09.36 s
p_hat300-1_c36-s279.clq	300	10933	8	2194	0	00	09.36 s
p_hat300-1_c36-s283.clq	300	10933	8	2194	0	00	09.38 s
p_hat300-1_c36-s287.clq	300	10933	8	2194	0	00	09.36 s
p_hat300-1_c54-s279.clq	300	10933	8	2194	0	00	09.35 s
p_hat300-1_c54-s283.clq	300	10933	8	2194	0	00	09.35 s
p_hat300-1_c54-s287.clq	300	10933	8	2194	0	00	09.36 s
p_hat300-2_c108-s291.clq	300	21928	25	6066	20	4590	15.52 s
p_hat300-2_c108-s295.clq	300	21928	25	6066	20	4590	15.51 s
p_hat300-2_c108-s299.clq	300	21928	25	6066	24	5479	15.49 s
p_hat300-2_c36-s291.clq	300	21928	25	6066	24	5479	15.50 s
p_hat300-2_c36-s295.clq	300	21928	25	6066	20	4590	15.44 s
p_hat300-2_c36-s299.clq	300	21928	25	6066	20	4590	15.44 s
p_hat300-2_c72-s291.clq	300	21928	25	6066	20	4590	15.50 s
p_hat300-2_c72-s295.clq	300	21928	25	6066	20	4590	15.56 s
p_hat300-2_c72-s299.clq	300	21928	25	6066	20	4590	15.56 s
p_hat300-3_c111-s303.clq	300	33390	33	7090	36	6606	23.28 s
p_hat300-3_c111-s307.clq	300	33390	34	7049	36	6506	23.33 s
p_hat300-3_c111-s311.clq	300	33390	33	7090	36	6288	23.38 s
p_hat300-3_c167-s303.clq	300	33390	33	7083	35	6494	23.33 s
p_hat300-3_c167-s307.clq	300	33390	34	7049	36	6506	23.33 s
p_hat300-3_c167-s311.clq	300	33390	33	7083	36	6751	23.33 s
p_hat300-3_c56-s303.clq	300	33390	33	7060	34	6163	15.49 s
p_hat300-3_c56-s307.clq	300	33390	33	7083	34	5889	23.33 s
p_hat300-3_c56-s311.clq	300	33390	34	7049	36	6506	00.00 s
MANN_a27_c188-s315.clq	378	70551	125	24476	126	24251	42.59 s
MANN_a27_c188-s319.clq	378	70551	125	24476	126	24251	43.39 s
MANN_a27_c188-s323.clq	378	70551	125	24476	126	24251	43.35 s
MANN_a27_c281-s315.clq	378	70551	125	24476	126	24251	43.01 s
MANN_a27_c281-s319.clq	378	70551	125	24476	126	24251	43.39 s
MANN_a27_c281-s323.clq	378	70551	125	24476	126	24251	43.11 s
MANN_a27_c94-s315.clq	378	70551	125	24476	126	24251	43.31 s
MANN_a27_c94-s319.clq	378	70551	125	24476	126	24251	43.21 s
MANN_a27_c94-s323.clq	378	70551	125	24476	126	24251	43.20 s
san400_0.5_1_c100-s327.clq	400	39900	13	3922	9	2623	23.00 s
san400_0.5_1_c100-s331.clq	400	39900	9	3146	9	2597	23.38 s

san400_0.5_1_c100-s335.clq	400	39900	8	3066	9	2888	23.14 s
san400_0.5_1_c150-s327.clq	400	39900	9	3093	9	2623	23.17 s
san400_0.5_1_c150-s331.clq	400	39900	8	3066	9	2687	00.00 s
san400_0.5_1_c150-s335.clq	400	39900	13	3922	10	2972	23.39 s
san400_0.5_1_c50-s327.clq	400	39900	13	3922	9	2878	23.00 s
san400_0.5_1_c50-s331.clq	400	39900	9	3106	9	3021	00.00 s
san400_0.5_1_c50-s335.clq	400	39900	8	3066	9	2687	23.38 s
sanr400_0.5_c100-s339.clq	400	39984	12	4168	13	3716	21.22 s
sanr400_0.5_c100-s343.clq	400	39984	12	4168	13	3706	21.56 s
sanr400_0.5_c100-s347.clq	400	39984	12	4168	13	3792	21.16 s
sanr400_0.5_c150-s339.clq	400	39984	12	4168	13	3706	21.22 s
sanr400_0.5_c150-s343.clq	400	39984	12	4168	13	3792	21.24 s
sanr400_0.5_c150-s347.clq	400	39984	12	4168	13	3716	22.42 s
sanr400_0.5_c50-s339.clq	400	39984	12	4036	13	3716	21.36 s
sanr400_0.5_c50-s343.clq	400	39984	12	4024	13	3716	21.18 s
sanr400_0.5_c50-s347.clq	400	39984	12	4168	13	3716	21.28 s
san400_0.7_1_c140-s351.clq	400	55860	21	7609	22	6180	33.01 s
san400_0.7_1_c140-s355.clq	400	55860	21	7634	22	6311	32.49 s
san400_0.7_1_c140-s359.clq	400	55860	40	10658	26	6977	32.57 s
san400_0.7_1_c210-s351.clq	400	55860	21	7493	22	6180	32.42 s
san400_0.7_1_c210-s355.clq	400	55860	21	7490	22	6180	32.51 s
san400_0.7_1_c210-s359.clq	400	55860	40	10658	27	7184	32.39 s
san400_0.7_1_c70-s351.clq	400	55860	21	7490	22	6044	33.27 s
san400_0.7_1_c70-s355.clq	400	55860	21	7580	22	6180	32.59 s
san400_0.7_1_c70-s359.clq	400	55860	21	7582	22	6180	32.46 s
san400_0.7_2_c140-s363.clq	400	55860	30	7974	23	6067	32.38 s
san400_0.7_2_c140-s367.clq	400	55860	30	7974	23	6067	32.39 s
san400_0.7_2_c140-s371.clq	400	55860	30	7974	23	6067	32.34 s
san400_0.7_2_c210-s363.clq	400	55860	30	7974	20	5170	32.44 s
san400_0.7_2_c210-s367.clq	400	55860	30	7974	23	6067	32.37 s
san400_0.7_2_c210-s371.clq	400	55860	30	7974	23	6067	32.34 s
san400_0.7_2_c70-s363.clq	400	55860	30	7974	23	6067	32.55 s
san400_0.7_2_c70-s367.clq	400	55860	30	7974	23	6067	32.46 s
san400_0.7_2_c70-s371.clq	400	55860	30	7974	22	6251	32.39 s
san400_0.7_3_c140-s375.clq	400	55860	22	5641	20	5188	34.31 s
san400_0.7_3_c140-s379.clq	400	55860	22	5641	20	4986	34.31 s

san400_0.7_3_c140-s383.clq	400	55860	22	5641	18	5156	34.21 s
san400_0.7_3_c210-s375.clq	400	55860	22	5641	20	4962	34.30 s
san400_0.7_3_c210-s379.clq	400	55860	22	5641	19	4757	34.32 s
san400_0.7_3_c210-s383.clq	400	55860	22	5641	19	4757	34.35 s
san400_0.7_3_c70-s375.clq	400	55860	22	5641	20	4962	34.26 s
san400_0.7_3_c70-s379.clq	400	55860	22	5641	17	4596	34.34 s
san400_0.7_3_c70-s383.clq	400	55860	22	5641	20	4962	34.27 s
sanr400_0.7_c140-s387.clq	400	55869	20	6472	21	5754	30.54 s
sanr400_0.7_c140-s391.clq	400	55869	20	6472	21	6043	30.45 s
sanr400_0.7_c140-s395.clq	400	55869	20	6472	21	6133	30.44 s
sanr400_0.7_c210-s387.clq	400	55869	20	6577	21	5704	31.00 s
sanr400_0.7_c210-s391.clq	400	55869	20	6472	21	6097	30.51 s
sanr400_0.7_c210-s395.clq	400	55869	20	6472	21	6175	30.48 s
sanr400_0.7_c70-s387.clq	400	55869	20	6536	21	6175	30.50 s
sanr400_0.7_c70-s391.clq	400	55869	20	6473	21	5704	30.46 s
sanr400_0.7_c70-s395.clq	400	55869	20	6577	21	5761	30.52 s
brock400_3_c148-s399.clq	400	59681	25	7253	24	6744	32.54 s
brock400_3_c148-s403.clq	400	59681	25	7253	25	6780	32.56 s
brock400_3_c148-s407.clq	400	59681	25	7253	23	6597	32.55 s
brock400_3_c222-s399.clq	400	59681	25	7253	24	6704	33.11 s
brock400_3_c222-s403.clq	400	59681	25	7253	24	6339	32.52 s
brock400_3_c222-s407.clq	400	59681	25	7253	24	6170	32.55 s
brock400_3_c74-s399.clq	400	59681	25	7253	24	6679	32.54 s
brock400_3_c74-s403.clq	400	59681	25	7253	24	5979	32.45 s
brock400_3_c74-s407.clq	400	59681	25	7253	24	6677	32.53 s
brock400_1_c148-s411.clq	400	59723	23	7063	24	6381	33.04 s
brock400_1_c148-s415.clq	400	59723	23	6791	24	6023	33.11 s
brock400_1_c148-s419.clq	400	59723	23	6795	24	5851	33.06 s
brock400_1_c222-s411.clq	400	59723	22	6822	23	6056	33.10 s
brock400_1_c222-s415.clq	400	59723	23	6791	24	6055	33.18 s
brock400_1_c222-s419.clq	400	59723	23	6950	24	5773	33.15 s
brock400_1_c74-s411.clq	400	59723	22	6852	23	5766	33.22 s
brock400_1_c74-s415.clq	400	59723	23	6791	24	5762	33.15 s
brock400_1_c74-s419.clq	400	59723	22	6882	23	5811	33.31 s
brock400_4_c148-s423.clq	400	59765	24	7056	25	6417	33.31 s
brock400_4_c148-s427.clq	400	59765	24	7134	25	6680	33.07 s

brock400_4_c148-s431.clq	400	59765	24	7069	25	6417	33.12 s
brock400_4_c222-s423.clq	400	59765	22	7095	23	5845	33.31 s
brock400_4_c222-s427.clq	400	59765	22	7196	23	5864	33.02 s
brock400_4_c222-s431.clq	400	59765	24	7077	25	6417	33.19 s
brock400_4_c74-s423.clq	400	59765	24	7077	25	6417	33.13 s
brock400_4_c74-s427.clq	400	59765	24	7077	25	6383	33.23 s
brock400_4_c74-s431.clq	400	59765	22	7095	24	6033	33.37 s
brock400_2_c148-s435.clq	400	59786	25	7280	24	5632	33.18 s
brock400_2_c148-s439.clq	400	59786	24	7163	25	5802	33.21 s
brock400_2_c148-s443.clq	400	59786	23	7119	24	6567	33.19 s
brock400_2_c222-s435.clq	400	59786	24	7090	25	6350	33.06 s
brock400_2_c222-s439.clq	400	59786	23	7123	24	5722	33.19 s
brock400_2_c222-s443.clq	400	59786	24	7163	24	6008	33.15 s
brock400_2_c74-s435.clq	400	59786	24	7090	25	6471	33.08 s
brock400_2_c74-s439.clq	400	59786	24	7090	25	5940	33.07 s
brock400_2_c74-s443.clq	400	59786	25	7280	23	6216	33.01 s
san400_0.9_1_c180-s447.clq	400	71820	100	21387	64	14879	42.36 s
san400_0.9_1_c180-s451.clq	400	71820	100	21387	65	15030	42.40 s
san400_0.9_1_c180-s455.clq	400	71820	100	21387	69	14757	42.43 s
san400_0.9_1_c270-s447.clq	400	71820	100	21387	64	14879	42.49 s
san400_0.9_1_c270-s451.clq	400	71820	100	21387	72	14556	42.43 s
san400_0.9_1_c270-s455.clq	400	71820	100	21387	65	14434	42.49 s
san400_0.9_1_c90-s447.clq	400	71820	53	15128	54	12132	43.30 s
san400_0.9_1_c90-s451.clq	400	71820	100	21387	70	14834	42.29 s
san400_0.9_1_c90-s455.clq	400	71820	100	21387	70	14649	42.33 s
johnson32-2-4_c108-s459.clq	496	107880	16	5577	0	00	74.11 s
johnson32-2-4_c108-s463.clq	496	107880	16	5577	0	00	74.16 s
johnson32-2-4_c108-s467.clq	496	107880	16	5577	0	00	74.12 s
johnson32-2-4_c216-s459.clq	496	107880	16	5577	0	00	74.46 s
johnson32-2-4_c216-s463.clq	496	107880	16	5577	0	00	74.31 s
johnson32-2-4_c216-s467.clq	496	107880	16	5577	0	00	74.26 s
johnson32-2-4_c324-s459.clq	496	107880	16	5577	0	00	74.42 s
johnson32-2-4_c324-s463.clq	496	107880	16	5577	0	00	74.31 s
johnson32-2-4_c324-s467.clq	496	107880	16	5577	0	00	74.23 s
c-fat500-1_c12-s471.clq	500	4459	14	6880	0	00	04.57 s
c-fat500-1_c12-s475.clq	500	4459	14	6880	0	00	04.56 s

c-fat500-1_c12-s479.clq	500	4459	14	6880	0	00	04.56 s
c-fat500-1_c4-s471.clq	500	4459	14	6880	0	00	04.56 s
c-fat500-1_c4-s475.clq	500	4459	14	6880	0	00	04.54 s
c-fat500-1_c4-s479.clq	500	4459	14	6880	0	00	04.56 s
c-fat500-1_c8-s471.clq	500	4459	14	6880	0	00	04.57 s
c-fat500-1_c8-s475.clq	500	4459	14	6880	0	00	04.56 s
c-fat500-1_c8-s479.clq	500	4459	14	6880	0	00	04.55 s
c-fat500-2_c18-s483.clq	500	9139	26	12544	0	00	06.16 s
c-fat500-2_c18-s487.clq	500	9139	26	12544	0	00	06.18 s
c-fat500-2_c18-s491.clq	500	9139	26	12544	0	00	06.19 s
c-fat500-2_c27-s483.clq	500	9139	26	12544	0	00	06.20 s
c-fat500-2_c27-s487.clq	500	9139	26	12544	0	00	06.18 s
c-fat500-2_c27-s491.clq	500	9139	26	12544	0	00	06.15 s
c-fat500-2_c9-s483.clq	500	9139	26	12544	0	00	06.27 s
c-fat500-2_c9-s487.clq	500	9139	26	12544	0	00	06.16 s
c-fat500-2_c9-s491.clq	500	9139	26	12544	0	00	06.20 s
c-fat500-5_c23-s495.clq	500	23191	64	29055	0	00	00.00 s
c-fat500-5_c23-s499.clq	500	23191	64	29055	0	00	14.58 s
c-fat500-5_c23-s503.clq	500	23191	64	29055	0	00	14.55 s
c-fat500-5_c45-s495.clq	500	23191	64	29055	0	00	14.55 s
c-fat500-5_c45-s499.clq	500	23191	64	29055	0	00	14.56 s
c-fat500-5_c45-s503.clq	500	23191	64	29055	0	00	14.42 s
c-fat500-5_c68-s495.clq	500	23191	64	29055	0	00	14.49 s
c-fat500-5_c68-s499.clq	500	23191	64	29055	0	00	15.18 s
c-fat500-5_c68-s503.clq	500	23191	64	29055	0	00	15.19 s
p_hat500-1_c32-s507.clq	500	31569	9	4145	0	00	30.20 s
p_hat500-1_c32-s511.clq	500	31569	9	4145	0	00	30.32 s
p_hat500-1_c32-s515.clq	500	31569	9	4145	0	00	44.00 s
p_hat500-1_c63-s507.clq	500	31569	9	4145	0	00	30.37 s
p_hat500-1_c63-s511.clq	500	31569	9	4145	0	00	30.32 s
p_hat500-1_c63-s515.clq	500	31569	9	4145	0	00	24.00 s
p_hat500-1_c94-s507.clq	500	31569	9	4145	0	00	38.00 s
p_hat500-1_c94-s511.clq	500	31569	9	4145	0	00	31.02 s
p_hat500-1_c94-s515.clq	500	31569	9	4145	9	3640	30.36 s
c-fat500-10_c139-s519.clq	500	46627	126	51344	0	00	30.30 s
c-fat500-10_c139-s523.clq	500	46627	126	51344	0	00	30.22 s

c-fat500-10_c139-s527.clq	500	46627	126	51344	0	00	30.18 s
c-fat500-10_c47-s519.clq	500	46627	126	51344	0	00	31.02 s
c-fat500-10_c47-s523.clq	500	46627	126	51344	0	00	30.36 s
c-fat500-10_c47-s527.clq	500	46627	126	51344	0	00	30.02 s
c-fat500-10_c93-s519.clq	500	46627	126	51344	0	00	30.36 s
c-fat500-10_c93-s523.clq	500	46627	126	51344	0	00	30.36 s
c-fat500-10_c93-s527.clq	500	46627	126	51344	0	00	30.36 s
p_hat500-2_c125-s531.clq	500	62946	35	13647	36	13485	55.15 s
p_hat500-2_c125-s535.clq	500	62946	35	13647	36	13343	55.27 s
p_hat500-2_c125-s539.clq	500	62946	35	13647	36	13511	56.14 s
p_hat500-2_c188-s531.clq	500	62946	35	13647	36	13324	56.05 s
p_hat500-2_c188-s535.clq	500	62946	35	13647	36	13511	55.49 s
p_hat500-2_c188-s539.clq	500	62946	35	13647	36	13116	56.15 s
p_hat500-2_c63-s531.clq	500	62946	35	13647	36	13229	55.29 s
p_hat500-2_c63-s535.clq	500	62946	35	13647	36	13471	55.29 s
p_hat500-2_c63-s539.clq	500	62946	35	13647	36	13485	55.10 s
p_hat500-3_c188-s543.clq	500	93800	48	16538	49	13772	86.18 s
p_hat500-3_c188-s547.clq	500	93800	48	16590	49	14498	86.08 s
p_hat500-3_c188-s551.clq	500	93800	48	16538	49	14720	86.36 s
p_hat500-3_c282-s543.clq	500	93800	48	16629	49	14189	86.16 s
p_hat500-3_c282-s547.clq	500	93800	48	16660	49	15150	85.40 s
p_hat500-3_c282-s551.clq	500	93800	48	16514	49	14240	86.09 s
p_hat500-3_c94-s543.clq	500	93800	48	16578	49	14422	85.34 s
p_hat500-3_c94-s547.clq	500	93800	48	16538	49	14051	85.21 s
p_hat500-3_c94-s551.clq	500	93800	48	16590	49	13631	85.04 s
p_hat700-1_c126-s555.clq	700	60999	11	6630	0	00	70.15 s
p_hat700-1_c126-s559.clq	700	60999	11	6630	0	00	68.45 s
p_hat700-1_c126-s563.clq	700	60999	11	6630	0	00	68.53 s
p_hat700-1_c42-s555.clq	700	60999	11	6630	0	00	69.23 s
p_hat700-1_c42-s559.clq	700	60999	11	6630	0	00	70.14 s
p_hat700-1_c42-s563.clq	700	60999	11	6630	0	00	68.58 s
p_hat700-1_c84-s555.clq	700	60999	11	6630	0	00	69.06 s
p_hat700-1_c84-s559.clq	700	60999	11	6630	0	00	69.14 s
p_hat700-1_c84-s563.clq	700	60999	11	6630	0	00	69.16 s
p_hat700-2_c172-s567.clq	700	121728	44	24230	44	24062	139.23 s
p_hat700-2_c172-s571.clq	700	121728	44	24230	44	24062	140.24 s

p_hat700-2_c172-s575.clq	700	121728	44	24230	44	23356	141.26 s
p_hat700-2_c258-s567.clq	700	121728	44	24230	44	23651	140.35 s
p_hat700-2_c258-s571.clq	700	121728	44	24230	43	23703	141.11 s
p_hat700-2_c258-s575.clq	700	121728	44	24230	44	24062	142.04 s
p_hat700-2_c86-s567.clq	700	121728	44	24230	44	24039	140.01 s
p_hat700-2_c86-s571.clq	700	121728	44	24230	44	24062	141.02 s
p_hat700-2_c86-s575.clq	700	121728	44	24230	44	24039	140.03 s
p_hat700-3_c130-s579.clq	700	183010	60	29860	62	27839	241.34 s
p_hat700-3_c130-s583.clq	700	183010	60	29728	61	27423	236.17 s
p_hat700-3_c130-s587.clq	700	183010	60	30055	61	27574	231.56 s
p_hat700-3_c259-s579.clq	700	183010	60	29792	61	28084	232.54 s
p_hat700-3_c259-s583.clq	700	183010	61	29727	62	28579	239.04 s
p_hat700-3_c259-s587.clq	700	183010	60	29753	61	27397	237.01 s
p_hat700-3_c389-s579.clq	700	183010	60	30025	62	28129	237.54 s
p_hat700-3_c389-s583.clq	700	183010	61	29664	62	28095	232.49 s
p_hat700-3_c389-s587.clq	700	183010	61	29727	62	27882	233.47 s