

POC サーバ: コミュニティ分析支援サーバ

1. はじめに

POC サーバは POC システムにおけるサーバシステムである。POC サーバの特徴は、POC システム実験過程において分析者のデータ収集・分析作業を支援する点である。以下、POC サーバの研究背景と実装システムの概要について述べる。

2. 研究背景

コミュニティ支援システムの開発サイクルについて述べ、コミュニティ分析支援機能の必要性と要件について述べる。

2.1 コミュニティ支援システムの開発サイクル

コミュニティ支援システムの開発では分析、設計、実験からなる開発サイクルが重要である。用途を限定しないコミュニティ支援システムは汎用的である反面、どのコミュニティにとっても不十分な機能しか提供できない危険性がある。従ってコミュニティ支援システムの開発では、実際のコミュニティの活動状況を観察し分析した上でシステム設計を行い、設計したシステムをコミュニティ内で評価し、評価結果をシステムに反映するといった開発サイクルが必要である。図 1 にコミュニティ支援システムの開発サイクルを示す。

コミュニティ支援システムの開発サイクルを円滑に進めるにはコミュニティ分析支援機能が必要である。コミュニティ分析支援機能とは、システムの実験過程において分析者のデータ収集・分析作業を支援する機能である。これまでコミュニティ支援システムの実証実験 [1, 2] や CMC(Computer-Mediated-Communication) 研究における実験 [3, 4] が報告されているが、実験データの収集と分析には人手を要し、実験の結果得られた知見を速やかにフィードバックすることは困難であった。またコミュニティ支援システムのフレームワークも提案されている [5, 6] が、分析者支援について考慮されていない。コミュニティ支援システムの円滑な開発サイクルの実現には、コミュニティ分析支援機能が必要である。

2.2 コミュニティ分析支援機能の要件

コミュニティ分析支援機能には次の要件が求められる。

1. データ収集・解析の自動化*1
2. 実験期間中のデータ分析支援
3. メッセージ分析作業の支援

第 1 に、データの収集と解析を自動化する必要がある。従来の実験過程ではデータの収集と分析に多くの時間を費やした。このため、コミュニティ分析支援機能は実験期間中、自動的にデータを収集し、分析者の求めに応じてデータを解析できる必要がある。ここではコミュニティ支援システムのサーバ上に記録されるサーバログと、コミュニティ参加者の発言であるメッセージを対象とする。

第 2 に、分析者は実験期間中もデータ分析できる必要がある。従来、収集したデータの分析には多くの時間を要するた

*1 本稿では解析と分析を使い分けている。解析は計算機による自動データ処理を指し、分析は人間の分析者によるデータ処理を指す。

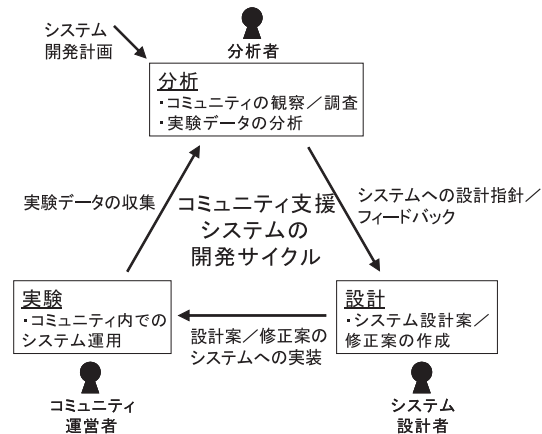


図 1: コミュニティ支援システム開発サイクル

め、実験期間中のリアルタイムな分析は困難であった。このため、コミュニティ分析支援機能は分析者の求めに応じ、その時点までの最新データを解析し分析者の作業を支援する必要がある。実装ではサーバログとメッセージ両方の最新データを扱えるようにする。

第 3 に、分析者のメッセージ分析作業を支援する必要がある。コミュニティにおいて、ある参加者の発言(メッセージ)は他の参加者のメッセージに影響を与える。このメッセージ間の関係を分析することで、分析者は意見の影響過程や話題の遷移過程を把握できる。しかしながら、従来では手作業による分析を余儀なくされ、例えば 1,000 件のメッセージがあった場合、メッセージ間の関係を分析するのは容易ではなかった。このことからコミュニティ分析支援機能は、大量のメッセージを迅速に解析できる必要がある。

3. POC サーバ

筆者らは POC サーバにおいて第 1 と第 2 の要件について実装した。第 3 の要件について POC Analyzer において実装した。ここでは POC サーバについて述べる。POC サーバにおける分析者支援機能は (1) コミュニティ運営モジュール、(2) メッセージ管理モジュール、(3) ログ分析モジュールにおいて実装されている。図 2 に POC システムにおける POC サーバの位置付けを示す。

3.1 POC サーバにおけるデータ収集方針

POC サーバにおけるデータ収集方針は次の通り。

サーバログを対象とする

サーバログには各 POC クライアントのメッセージ閲覧や投稿行動のログが記録されている。このため分析者はサーバログを分析することで、コミュニティ内で中心的

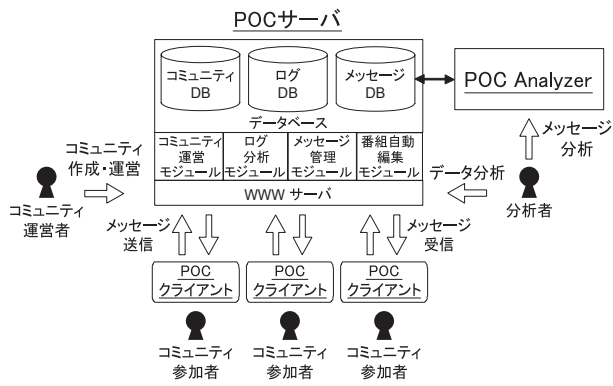


図 2: POC システムにおける POC サーバの位置付け。

表 1: トークン一覧

トークン	説明
TRY LOGIN	コミュニティへのログイン要請
ACK LOGIN	ログイン成功
NACK LOGIN	ログイン失敗
LIST COMMUNITIES	コミュニティー一覧取得
LIST MESSAGES	メッセージ一覧取得
SEARCH MESSAGES	メッセージ検索
ACK MESSAGE	メッセージ登録
DELETE MESSAGE	メッセージ削除
LOGOUT	コミュニティからのログアウト

役割を果たす人物を把握したり、円滑な運営の行われているコミュニティに関する知見を得られる可能性がある。

ログフォーマットを統一する

POC システムには複数の POC クライアントが存在するが、それぞれが独自のログフォーマットを用いては分析の手間である。このため POC サーバは表 1 のトークンを使ってログフォーマットを統一し、ログ解析処理を一元化する。表 2 にトークンを使ったログを示す。表 2 は 'person1' のログインからログアウトまでの行動を示している。

3.2 コミュニティ運営モジュール

コミュニティ運営モジュールはコミュニティの設定や参加者のアカウント管理等を行うためのモジュールである。図 3 にコミュニティ運営モジュールの提供するコミュニティー一覧及び設定画面を示す。運営者は各コミュニティに投稿されたメッセージ数や最後にメッセージの投稿された日時等の情報を把握できる。また各コミュニティに所属する参加者のアカウントを登録・変更・削除できる。

3.3 メッセージ管理モジュール

メッセージ管理モジュールは各コミュニティのメッセージとストーリーを管理・閲覧するためのモジュールである。運営者は各コミュニティのメッセージとストーリーの内容を閲覧できる。図 4 にストーリー内容表示画面を示す。運営者は Web

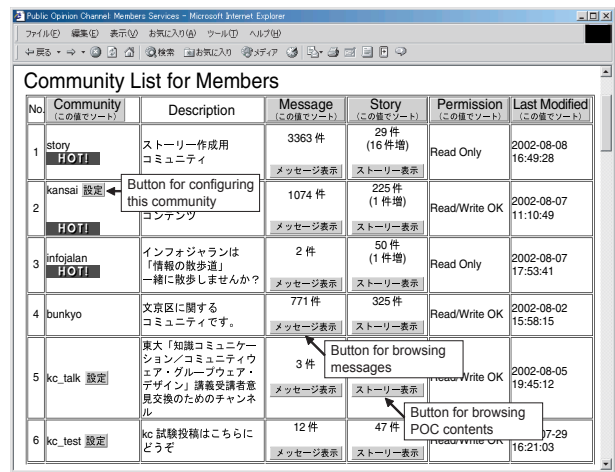


図 3: コミュニティー一覧と設定画面。

ブラウザ上でメッセージとストーリーの検索・閲覧・削除を行える。

3.4 ログ解析モジュール

ログ解析モジュールはサーバログを解析し、解析結果をレポートとして分析者に提供する。図 5 にレポートの例を示す。図 5 はあるコミュニティへのアクセス状況を示しており、アクセス頻度のグラフと共に実験期間中の平均アクセス数や最長・最短アクセス時間等の統計情報が記載されている。分析者はレポートから実験期間中のコミュニティの状態を容易に把握できる。

ログ解析モジュールの特徴は次の通り。

1. 最新のログを用いた解析
2. 分析目的に応じた柔軟なログ分析支援

第 1 に、ログ解析モジュールは最新ログを用いた解析を行える。分析者は最新ログに基づき実験期間中、コミュニティの現状を的確に判断できる。

第 2 に、ログ解析モジュールは分析者の目的に応じた柔軟な分析を支援する。分析者は分析範囲を指定し任意の期間のログを分析できる。またログ解析モジュールは次に示す様々なレポートを分析者に提供する。

- 1 日当たりのログイン数及びメッセージ投稿数
- コミュニティ参加者毎のセッション情報 (ログインからログアウトまでの接続時間と接続先コミュニティ)
- コミュニティ別の総投稿数
- コミュニティ別及び時間帯別の投稿数の内訳
- 過去 3 日間に投稿されたメッセージ一覧

これらのレポートから分析者は、コミュニティ参加者の立ち寄ったコミュニティと滞在時間、滞在中の行動といった参加者単位の分析から、あるコミュニティにおけるアクセス数、コミュニティ間での投稿数の比較といったコミュニティ単位の分析まで、様々な視点でログ分析できる。図 6 に各種レポートのグラフを示す。

4. 評価と考察

FTTH トライアルにおける POC サーバの効果について述べる。

表 2: サーバログの例

行	イベント
1	TRY LOGIN person=person1 community=kc_talk [2002-07-10 19:09:55]
2	ACK LOGIN person=person1 community=kc_talk [2002-07-10 19:09:56]
3	LIST COMMUNITIES IP=xxx.xxx.xxx.xx person=person1 clientName=COMMUNICATOR [2002-07-10 19:09:58]
4	LIST MESSAGES client=COMMUNICATOR person=person1 [2002-07-10 19:10:00]
5	SEARCH MESSAGE query='風' clientName=COMMUNICATOR community=kc_talk found=47 [2002-07-10 19:16:20]
6	ACK MESSAGE IP=xxx.xxx.xxx.xx file=bbs107.xml person=person1 community=kc_talk [2002-07-10 19:23:31]
7	LOGOUT community=kc_talk IP=xxx.xxx.xxx.xx person=person1 [2002-07-10 19:48:43]

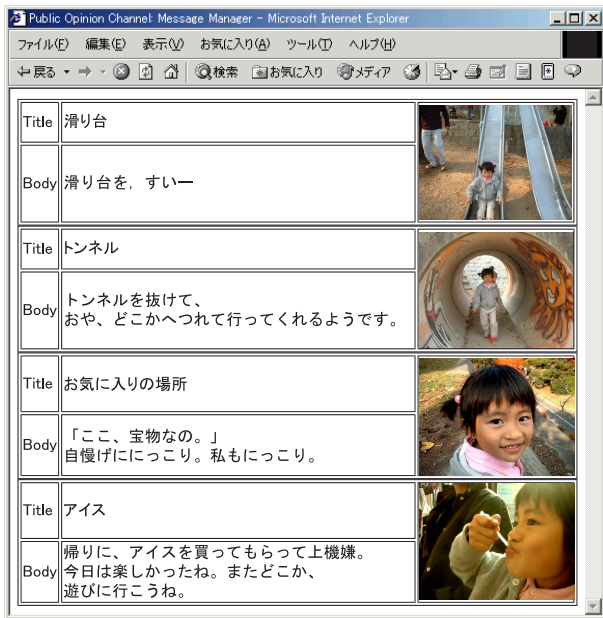


図 4: ストーリー内容表示画面

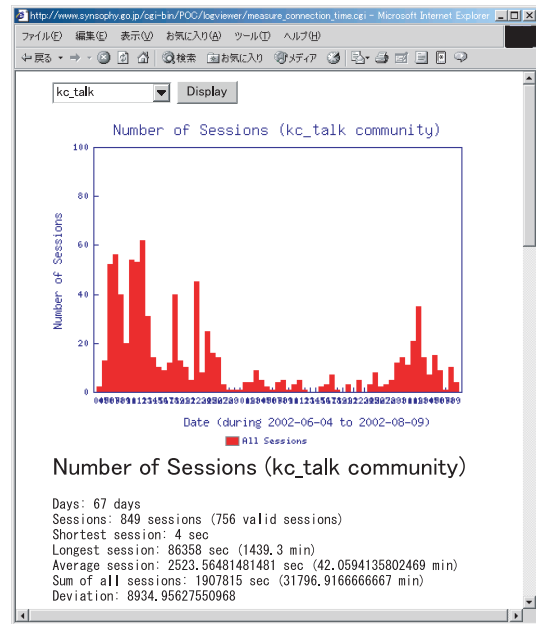


図 5: ログ解析モジュールによるレポートの例

4.1 評価実験: FTTH トライアル

筆者らは 2002 年 3 月 25 日から 9 月 30 日までの間、東京都内の 444 世帯を対象としたブロードバンドネットワークの実証実験 (KDDI FTTH トライアル) において POC システムの評価実験を行った。実験期間中のサーバログは 68,000 行 (6.3MByte) であった。この事例ではログ解析モジュールについて次の効果を確認した。

ログ解析モジュールは分析者の実験状況把握を支援した。実験期間中、分析者はアクセス数やメッセージ投稿数といったコミュニティの現状を知りたいという要求を持っていた。しかしながら従来の分析過程では、分析者自らがログを収集し、データの整形や不要なデータの削除といった前処理を施してから Microsoft Excel^{*2} や SPSS^{*3} といった表計算ソフトや統計パッケージを用いて分析しなければならなかった。特にデータの前処理は分析者にとって手間のかかる作業であり、コミュニティの現状を把握するまでに多くの時間を要した。

ログ解析モジュールは実験期間中の最新ログを解析できるた

め、半年間に及ぶ実験期間中にあっても分析者は常に最新のコミュニティの状況を把握できた。また分析者は実験開始当初のログと現在のログを比較したり、ある期間のログを取り出して分析するといった分析作業を柔軟に行えた。

4.2 考察

表 3 に従来の分析方法と POC サーバによる分析方法の比較を示す。POC サーバにより大量のデータに対して迅速な分析作業が可能になると言える。また実験期間中のデータ分析と探索的なデータ分析に関して、POC サーバの利用により迅速で柔軟なデータ分析作業が可能になると言える。

次に心理学実験における POC サーバの効果について述べる。心理学実験はコミュニティ支援システムの効果の測定に重要である。POC サーバは次の点で心理学実験を寄与し得る。

実験状況へのアウェアネスの獲得

実験者は実験状況についての気付き (アウェアネス) を獲得できる。従来、実験者は実験期間中のコミュニティの状態 (実験状況) を容易に把握できなかった。このため実験期間中、コミュニティに生じる興味深い現象を見過ご

*2 <http://www.microsoft.com/office/excel/>

*3 <http://www.spss.com/>

表 3: POC サーバの効果
手作業による分析

	手作業による分析	POC サーバ
データ収集に要する期間	2~3日	不要 ¹
扱い得るデータ量	100 メッセージ (0.4MByte)	68,000 行 (6.3MByte)
データ分析に要する時間	1 週間程度	10~20 秒 ²
実験期間中のデータ分析	可能だが時間を要する	最新データによる分析が可能
探索的なデータ分析	柔軟な分析が可能だが時間を要する	様々な条件の下で可能

¹ データは全てデータベースに格納されるため。

² CPU: Pentium II 333MHz, メモリ: 256MB において。

³ CPU: Pentium II 400MHz, メモリ: 512MB において。

のデータとの比較作業は手作業で処理しなければならない。今後、POC サーバの扱うログ形式を拡張し、他のコミュニケーションツールのデータと POC のデータを容易に比較できるよう改善する必要がある。

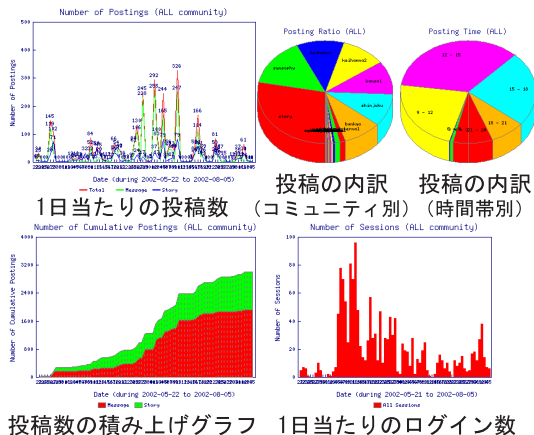
5. まとめ

POC サーバについて述べた。コミュニティ支援システムの開発にはコミュニティ分析支援機能が必要である。筆者らはコミュニティ分析支援機能を POC サーバに実装し、その有効性を確認した。

POC サーバの開発に当たっては、実際の実験に携わる心理学者からの助言が非常に有益であった。これまでに実施した実験からは POC サーバに対する多くの要求が生じ、その度にシステムを改善してきた経緯がある。本稿で述べたコミュニティ支援システムの開発サイクルにもあるよう、心理学者のコミュニティから徐々にフィードバックを得ながら POC サーバの改善を行う。

参考文献

- [1] 石田亨, 西村俊和, 八槇博史, 後藤忠弘, 西部喜康, 和氣弘明, 森原一郎, 服部文夫, 西田豊明, 武田英明, 沢田篤史, 前田晴美. モバイルコンピューティングによる国際会議支援. 情報処理学会論文誌, Vol. 39, No. 10, pp. 2855-2865, Oct 1998.
- [2] 松田晃一, 三宅貴浩. パーソナルエージェント指向仮想社会 PAW (第 2 版) の構築と評価. 情報処理学会論文誌, Vol. 41, No. 10, pp. 2698-2707, Oct 2000.
- [3] 篠原一光, 三浦麻子. WWW 掲示板を用いた電子コミュニティ形成過程に関する研究. 社会心理学研究, Vol. 14, No. 3, pp. 144-154, 1999.
- [4] 三浦麻子, 篠原一光. CMC における状況の認知と情報発信行動. 応用心理学研究, Vol. 27, No. 1, pp. 25-35, 2001.
- [5] M. Koch and M.S. Lacher. Integrating community services: A common infrastructure proposal. In *Knowledge-Based Intelligence Engineering Systems & Allied Technologies (KES'2000)*, pp. 56-59, 2000.
- [6] 吉田仙, 亀井剛次, 大黒毅, 桑原和宏. ネットワークコミュニティ支援システムのエージェント指向フレームワーク Shine. 情報処理学会論文誌, Vol. 43, No. 2, pp. 499-512, Feb 2002.



投稿数の積み上げグラフ 1日当たりのログイン数

図 6: ログ解析モジュールの提供する各種グラフ。

してしまう問題が存在した。これに対し POC サーバは実験者に最新の実験状況を提供するため、実験者は興味深い現象に気付きやすくなる。

実験操作タイミングの把握

実験者は適切な実験操作のタイミングを判断できる。従来、分析者は実際のコミュニティの発言状況に基づき、実験補助員(サクラ)や刺激となるメッセージの投入といった実験操作のタイミングを経験的に判断しなけりななかつた。一方、本提案により実験者は実験状況を的確に判断できるようになり、より適切な実験操作のタイミングを判断できる。

4.3 課題

POC サーバの課題として次の問題点が指摘された。

クライアントログへの対応

ログ解析モジュールはクライアントログを扱っていない。POC クライアントのマウスイベントやキー入力といったクライアントログの分析により、コミュニティ参加者のより詳細な行動分析を期待できる。クライアントログの収集と分析についても検討する必要がある。

解析対象データ

POC システムのデータしか扱えない。現在の POC サーバは POC のサーバログしか扱えないため、メーリングリストや電子掲示板等の他のコミュニケーションツール