

Zuverlässiger Endsystem-Multicast in MANETs

Steffen Blödt
Institut für Telematik — Universität Karlsruhe (TH)
bloedt@tm.uka.de

Abstract: Zuverlässige Multicast-Kommunikation spielt für viele Anwendungen eine wichtige Rolle. In dieser Arbeit wird das Endsystem-Multicast-Protokoll Narada, das ursprünglich für das feste Internet entworfen wurde, für den Einsatz in MANETs angepasst. Dazu gehört die Berücksichtigung des gemeinsam genutzten Mediums, eine intelligente Suche von Overlay-Nachbarn und eine verbesserte Metrik zur Overlay-Optimierung. Darüber hinaus wird das Protokoll um Mechanismen zur Fehlerbehebung erweitert.

1 Einleitung

Zuverlässige Multicast-Kommunikation spielt bei Anwendungen wie Whiteboard, Chat oder Multi-Player-Spielen eine wichtige Rolle. Herkömmliche Multicast-Transportprotokolle aus dem festen Internet sind in mobilen Ad-hoc-Netzen (MANET) jedoch nicht einsetzbar, da sie oft auf zentralen Komponenten basieren oder die Gegebenheiten in MANETs nicht berücksichtigen. In MANETs bieten sich vor allem Peer-to-Peer-basierte Endsystem-Multicast-Protokolle an, da beide viele Gemeinsamkeiten besitzen. Sowohl MANETs, als auch Peer-to-Peer-Anwendungen arbeiten dezentral, selbstorganisierend und alle beteiligten Komponenten sind gleichzeitig Client, Server und Router. Endsystem-Multicast-Protokolle abstrahieren von der darunterliegenden Netztopologie und verteilen die Daten über ein Overlay-Netz, das oberhalb der Netzwerkschicht die Gruppenmitglieder durch Transportverbindungen miteinander verbindet. Darüber hinaus wird auf Netzwerkschicht kein Multicast-Routing benötigt, was den einfachen Einsatz ermöglicht, da nur die Gruppenmitglieder das Endsystem-Multicast-Protokoll beherrschen müssen.

In diesem Paper wird das Endsystem-Multicast-Protokoll Narada [CRZ00], das ursprünglich für den Festnetzeinsatz entworfen wurde, für den Einsatz in MANETs angepasst. Dazu gehört die Berücksichtigung des geteilten Mediums und eine verbesserte Anpassung an die sich verändernde Netztopologie. Darüber hinaus wird das Protokoll um Mechanismen zur Fehlerbehebung erweitert. Dies geschieht durch den Einsatz von TCP für die Overlay-Verbindungen. Da hierdurch nur Punkt-zu-Punkt-Verbindungen abgesichert werden, wird das Protokoll um einen Mechanismus zur globalen Fehlerbehebung erweitert.

2 Bisherige und verwandte Arbeiten

Die beiden zuverlässigen Multicast-Transportprotokolle RALM [TOLG02] und ReACT [ROY⁺04] wurden für den Einsatz in MANETs entworfen und setzen auf einem unzu-

verlässigen Multicast-Verteildienst auf Netzwerkschicht auf. Dies hat den Nachteil, dass alle Knoten diesen Multicast-Verteildienst unterstützen müssen. Bei Endsystem-Multicast-Protokollen müssen nur die Gruppenmitglieder das entsprechende Protokoll beherrschen.

Unzuverlässige Endsystem-Multicast-Protokolle aus dem festen Internet sind beispielsweise NICE [BBK02], CAN [RHKS01] und Bayeux [ZZJ⁺01]. Diese haben alle den Nachteil, dass sie entweder auf zentralen Komponenten basieren oder ein speziell strukturiertes Overlay-Netz benötigen, das in MANETs unter Umständen schwer zu warten ist. Aus diesem Grund wird in dieser Arbeit das Narada-Protokoll [CRZ00] als Basis verwendet. Es benutzt wie Gossamer [CMB00] ein unstrukturiertes Overlay-Netz, das leicht zu warten und sehr robust ist. Außerdem besitzt Narada einen Mechanismus zur Anpassung des Overlays an die Netztopologie, was in MANETs sehr wichtig ist.

2.1 Das Narada-Protokoll

Das Narada-Protokoll arbeitet zweistufig. Zuerst wird ein robustes, unstrukturiertes Overlay aufgebaut; danach werden mit Hilfe eines modifizierten DVMRP-Protokolls quellen-spezifische Verteilbäume zur unzuverlässigen Datenverteilung auf dem Overlay etabliert.

Bei Narada hat jedes Gruppenmitglied eine vollständige Liste aller Gruppenmitglieder. Dies wird durch die Verbreitung von *Refresh*-Nachrichten zwischen den Overlay-Nachbarn erreicht. Diese enthalten jeweils eine aktuelle Liste der dem Knoten bekannten Gruppenmitglieder. Durch Sequenznummern wird sichergestellt, dass alte Einträge von neueren verdrängt werden. Diese Nachrichten ermöglichen auch die Erkennung von Overlay-Partitionierungen. In diesem Fall wird versucht das Overlay durch Verbindungen zur anderen Partition zu reparieren.

Ein wichtiger Teil von Narada ist die permanente Overlay-Optimierung. Dies ist vor allem in MANETs wichtig, da hiervon die Effizienz des Overlays abhängt. Dazu überprüft jeder Knoten periodisch die Distanz zu einem zufälligen anderen Gruppenmitglied mit Hilfe einer *Round-Trip-Time*-(RTT)-Messung. Daraus wird die Nützlichkeit einer zusätzlichen Overlay-Verbindung zu diesem Knoten berechnet. Die Nützlichkeit einer Verbindung ist die Summe der relativen Latenzzeit-Verbesserungen zu allen anderen Gruppenmitgliedern durch diese Verbindung. Falls die Nützlichkeit t_{add} einen gewissen Schwellenwert übersteigt, wird die Verbindung aufgebaut. Falls die Nützlichkeit einer bestehenden Verbindung unter einen Schwellenwert t_{drop} sinkt, wird die Verbindung wieder abgebaut. Um Oszillationen im Overlay zu verhindern gilt $t_{add} > t_{drop}$.

3 Das Narada-MAN-Protokoll

In diesem Abschnitt wird das Narada-MAN-Protokoll vorgestellt, das Narada für den Einsatz in MANETs optimiert. Zusätzlich besitzt Narada-MAN Mechanismen zur Fehlerbehebung, um eine zuverlässige Datenübertragung zu gewährleisten.

3.1 Ausnutzung des gemeinsam genutzten Broadcast-Mediums

Bei Endsystem-Multicast-Protokollen sind *alle* Gruppenmitglieder in einem Overlay-Netz aus Unicast-Transportverbindungen miteinander verbunden. Dies ist gerade in MANETs sehr ineffizient, da sich hier oft mehrere Knoten in unmittelbarer Übertragungsreichweite voneinander befinden. Ein Overlay-Netz zwischen diesen Knoten führt zu einer unnötigen Netzbelastung, da sich alle Knoten zwar in einem einzigen Hop erreichen könnten, aber Umwege über das Overlay nehmen müssen. In Abbildung 1 ist beispielhaft der Ausschnitt eines Overlay-Netzes dargestellt. Hier sendet Knoten *S* Multicast-Daten zu den Knoten *A*, *B*, *C* und *D*, die bis auf *D* alle in direkter Übertragungsreichweite liegen.

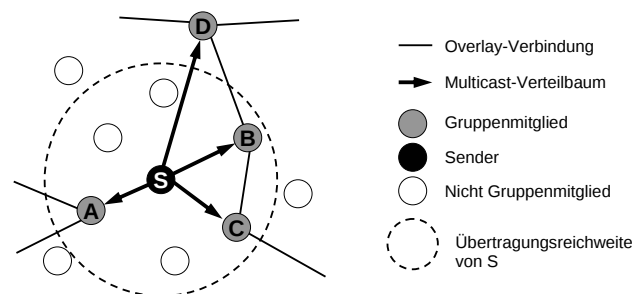


Abbildung 1: Ausnutzung des Broadcast-Mediums

Um dieses Problem zu beheben, kann jeder Knoten über die Einträge in seiner Routingtabelle auf Netzwerkschicht erfahren, welche Overlay-Nachbarn sich in direkter Übertragungsreichweite befinden. Verteilt ein Knoten Multicast-Daten an mehrere andere Knoten und sind mindestens zwei dieser Knoten in direkter Übertragungsreichweite, so versendet er an diese Knoten ein einziges Broadcast-Paket, anstatt mehrere einzelne Unicast-Pakete. Die weiter entfernten Knoten werden weiterhin durch Unicast-Pakete mit Daten versorgt.

Da bei diesem Verfahren Informationen des Routingprotokolls auf der Netzwerkschicht verwendet werden, spielt das verwendete Routingprotokoll eine wichtige Rolle. Bei proaktiven Protokollen wie OLSR sind diese Informationen immer aktuell. Bei reaktiven Protokollen wie AODV kann es jedoch passieren, dass zeitweise gar keine Informationen verfügbar sind. Dies ist bei Narada allerdings unbedenklich, da zwischen Overlay-Nachbarn sowieso periodisch Kontrollinformationen ausgetauscht werden, die dafür sorgen, dass die Routingtabellen immer gültige Werte enthalten.

3.2 Verbesserte Optimierungsmetrik

Narada verwendet zur Optimierung des Overlays eine Delay-Metrik mit RTT-Messungen. Aufgrund von Verzögerungen beim Medienzugriff und durch Übertragungswiederholungen auf der MAC-Schicht sind diese Messungen aber oft sehr unzuverlässig. Aus diesem Grund verwendet Narada-MAN eine kombinierte Hop/Delay-Metrik. Die RTT wird wie bei Narada durch Ping-Nachrichten bestimmt; die Hop-Distanz kann aus der Routingta-

belle auf Netzwerkschicht abgeleitet werden. Die Einträge dort sind immer gültig, da diese zuvor durch die Ping-Nachricht aufgefrischt wurden. Bei Narada-MAN werden beide Werte zur Bestimmung der Nützlichkeit einer Verbindung jeweils gleich gewichtet. In die Berechnung der Nützlichkeit gehen somit die relative Verbesserung der Hop-Anzahl und der RTT zu gleichen Teilen ein.

3.3 Zuverlässige Multicast-Übertragung

Um über das Overlay eine zuverlässige Datenübertragung zu gewährleisten, werden bei Narada-MAN zunächst die einzelnen Overlay-Verbindungen abgesichert, indem das zuverlässige Unicast-Transportprotokoll TCP verwendet wird. Dies übernimmt darüber hinaus zusätzlich Aufgaben wie Flusssteuerung und Staukontrolle. Durch diese Maßnahme werden allerdings nur die einzelnen Punkt-zu-Punkt-Verbindungen im Overlay abgesichert. Durch Umkonfigurationen der Overlay-Topologie und einer noch nicht erfolgten Anpassung des Datenverteilsbaumes kann es kurzfristig zu einem Paketverlust kommen. Hinzu kommen Paketverluste durch die verwendeten Broadcast-Übertragungen, bei denen kein TCP eingesetzt wird. Um dieses Problem zu lösen wird ein weiterer Mechanismus zur Fehlerbehebung benötigt. Zu diesem Zweck werden selektive Übertragungswiederholungen mit negativen Quittungen (NACK) eingesetzt. Falls ein Knoten einen Paketverlust aufgrund fehlender Sequenznummern erkennt, sendet er an seinen Vorgänger im Verteilsbaum ein NACK-Paket mit den fehlenden Sequenznummern. Jeder Knoten hat dazu einen Ringpuffer mit den zuletzt empfangenen Paketen. Falls der NACK-Empfänger die angeforderten Pakete nicht besitzt, sendet er das NACK-Paket rekursiv bis zum Sender weiter. NACKs und die dazugehörigen Paketwiederholungen werden dabei nicht gebroadcastet, sondern immer über die gesicherten TCP-Verbindungen versendet. Durch diesen Mechanismus kann theoretisch keine absolut zuverlässige Übertragung garantiert werden, da der verwendete Ringpuffer überlaufen kann und somit Pakete verloren gehen können. Bei entsprechender Dimensionierung des Puffers ist dies allerdings sehr unwahrscheinlich. Außerdem muss dieser Mechanismus nur die Paketverluste beheben, die durch eher seltene Umkonfigurationen im Overlay zustandekommen.

3.4 Intelligente Suche von Overlay-Nachbarn

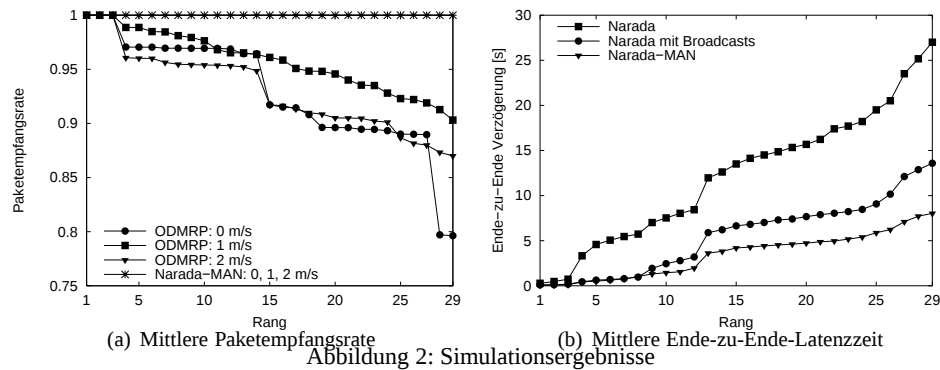
Für Endsystem-Multicast-Protokolle ist es sehr wichtig ein effizientes Overlay aufzubauen, um die Latenzzeit zu minimieren und die Netzressourcen zu schonen. Narada überprüft dazu periodisch zufällig ausgewählte Gruppenmitglieder, ob eine Verbindung zu ihnen für die Effizienz des Overlays nützlich ist. Da gerade beim Einsatz von TCP in MANETs lange Verbindungen sehr ungünstig sind, werden bei Narada-MAN bevorzugt nahe Gruppenmitglieder geprüft. Dazu werden Knoten im Abstand von einem, zwei, drei und mehr Hops mit einer Wahrscheinlichkeit von p_1 , p_2 , p_3 und p_o geprüft. Dabei gilt $p_1 > p_2 > p_3 > p_o$ und $\sum p_i = 1$. Dies führt dazu, dass nahe Knoten beim Aufbau neuer Verbindungen bevorzugt werden. Die Hop-Distanz wird dabei aus der Routingtabelle abgeleitet.

4 Evaluation

In diesem Abschnitt sind einige Simulationsergebnisse des Simulators GloMoSim dargestellt. Es wurden 80 Knoten mit einer Übertragungsbereichweite von 150 m in einem Gebiet von $1000 \text{ m} \times 1000 \text{ m}$ simuliert. 30 der Knoten sind Mitglieder der Multicast-Gruppe, wobei ein einzelner Sender nach 180 s mit einer konstanten Datenrate von 1 kByte/s Pakete sendet. Es wurde IEEE 802.11 mit einer Bandbreite von 2 Mbit/s und AODV als Routing-Protokoll verwendet. Als Mobilitätsmodell kam das *Reference Point Group Mobility Model* (RPGM) zum Einsatz.

Abbildung 2(a) vergleicht die Paketempfangsrate von Narada-MAN mit dem unzuverlässigen ODMRP bei unterschiedlicher Knotengeschwindigkeit. Die Empfangsrate ist dabei über dem Knotenrang aufgetragen. Dazu werden die Knoten nach jeder Simulation nach absteigender Empfangsrate sortiert und schließlich über alle Simulationen der Mittelwert über die Knoten mit gleichem Rang gebildet. Narada-MAN überträgt in allen Fällen die Pakete zuverlässig, weshalb hier nur eine Kurve zu sehen ist. Selbst bei einer Geschwindigkeit von 2 m/s, die häufige Umkonfigurationen des Overlays erfordert, können alle Paketverluste behoben werden. Bei ODMRP hingegen gehen bis zu 20% der Pakete verloren.

Abbildung 2(b) zeigt für $v_{max} = 1 \text{ m/s}$ die mittlere Ende-zu-Ende-Verzögerung für das herkömmliche Narada-Protokoll, für Narada mit Ausnutzung des Broadcast-Mediums und für Narada-MAN mit kombinierter Hop/Delay-Metrik. Man sieht, dass durch die vorgestellten Konzepte eine erhebliche Verbesserung erzielt wurde.



5 Schluss

In diesem Paper wurde das Narada-Protokoll für den Einsatz in MANETs verbessert und um Mechanismen zur Fehlerbehebung erweitert. Die Optimierungen bezogen sich auf die Ausnutzung des geteilten Mediums, eine intelligente Suche von Overlay-Nachbarn und eine verbesserte Optimierungsmetrik. Simulationen haben die Wirksamkeit der vorgestellten Verbesserungen gezeigt. Leider führt der Einsatz von TCP in MANETs zu erhebli-

chen Problemen durch hohe Latenzzeiten. Die Effizienz von Narada-MAN könnte deshalb durch eine speziell angepasstes Transportprotokoll weiter verbessert werden.

An dieser Stelle möchte ich Eckehardt Luhm für die Mitarbeit danken. Diese Arbeit wurde von der Deutschen Forschungsgemeinschaft (DFG) unterstützt.

Literatur

- [BBK02] Banerjee, S., Bhattacharjee, B., und Kommareddy, C.: Scalable application layer multicast. In: *ACM SIGCOMM 2002*. Pittsburgh. 2002.
- [CMB00] Chawathe, Y., McCanne, S., und Brewer, E. An architecture for Internet content distribution as an infrastructure service. Unpublished work. 2000.
- [CRZ00] Chu, Y., Rao, S. G., und Zhang, H.: A case for end system multicast. In: *ACM SIGMETRICS 2000*. S. 1–12. Santa Clara. 2000.
- [RHKS01] Ratnasamy, S., Handley, M., Karp, R., und Shenker, S.: Application-level multicast using content-addressable networks. *Lecture Notes in Computer Science*. 2001.
- [ROY⁺04] Rajendran, V., Obraczka, K., Yi, Y., Lee, S.-J., Tang, K., und Gerla, M.: Combining source- and localized recovery to achieve reliable multicast in multi-hop ad hoc networks. In: *Networking 2004*. May 2004.
- [TOLG02] Tang, K., Obraczka, K., Lee, S.-J., und Gerla, M.: A reliable, congestion-controlled multicast transport protocol in multimedia multi-hop networks. In: *ISCC 2002*. Taormina, Italy. July 2002.
- [ZZJ⁺01] Zhuang, S. Q., Zhao, B. Y., Joseph, A. D., Katz, R. H., und Kubiawicz, J. D.: Bayeux: An architecture for scalable and fault-tolerant wide-area data dissemination. In: *International Workshop on Network and Operating System Support for Digital Audio and Video*. June 2001.